

78622

STIC-Biotech/ChemLib

From: Davis, Minh-Tam
Sent: Thursday, October 24, 2002 12:45 PM
To: STIC-Biotech/ChemLib
Subject: Search request for 09/807200

Please search in commercial data base and in issued patent files:

SEQ ID NO:1

Thank you.

MINH TAM DAVIS

ART UNIT 1642, ROOM 8A01, MB 8E12

305-2008

CRFE

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

10/31/02

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2002, 06:21:57 ; Search time 1987 Seconds
(without alignments)
12498.439 Million cell updates/sec

Title: US-09-807-200-1

Perfect score: 1840

Sequence: 1 ctcgagccgggtggtggcag.....ctgttgctcaaaaaaaaaa 1840

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estnu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	990.8	53.8	1025	9	AL574221
2	981.2	53.3	1009	9	AL554923
3	949.6	51.6	963	9	AL546511
C 4	924	50.2	989	9	AL564981
5	909	49.4	971	9	AL550317
C 6	899	48.9	1011	9	AL574979
C 7	893.4	48.6	981	9	AL575679
C 8	891.6	48.5	950	9	AL577524
9	871.4	47.4	935	9	AL552820
10	870	47.3	934	9	AL552365
11	838.8	45.6	904	9	AL551401
C 12	836.2	45.4	854	9	AL577445
13	832.2	45.2	843	10	BG741018
14	831.6	45.2	979	9	AL551786
15	828.8	45.0	904	9	AL542564
16	824.2	44.8	931	9	AL547912
17	805.4	43.8	838	9	AL551123

18 792.8 43.1 869 10 BI759705
c 19 792.6 43.1 865 9 AL569820
20 786.2 42.7 881 10 BI761781
21 777.4 42.2 801 9 AL554806
22 752.6 40.9 825 9 AL543187
23 750 40.8 850 10 BI907864
24 742.6 40.4 795 10 BI755354
c 25 742.6 40.3 987 9 AL571835
26 741.6 40.3 803 10 BI757817
c 27 741.6 40.3 804 9 AL572678
28 736.6 40.0 852 10 BI764770
29 732.8 39.8 902 10 BI760791
30 723.2 39.3 783 10 BI752958
31 722.4 39.3 736 10 BI768850
32 721.4 39.2 863 10 BI912586
33 708.6 38.5 821 10 BI908447
c 34 697.2 37.9 831 9 AL542563
35 681.4 37.0 683 10 BI753249
36 678.8 36.9 842 10 BI819179
37 676 36.7 738 10 BI909271
38 671.8 36.5 711 10 BI905973
39 643.2 35.0 730 10 BI763703
c 40 628 34.1 638 10 BF435509
41 627.2 34.1 667 10 BI753288
42 627.2 34.1 682 10 BE407510
c 43 625.6 34.0 741 9 AL575359
c 44 624.4 33.9 629 9 AL574780
45 624.4 33.9 850 10 BI765797

ALIGNMENTS

RESULT 1
AL574221/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

AL574221 1025 bp mRNA linear EST 16-FEB-2001
AL574221 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1039YK16 3
prime. mRNA sequence.

AL574221 GI:12934220

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .1025

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1039YK16"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; site.1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@life.com URL :

http://fulllength.invitrogen.com"

192 a 306 c 343 g 176 t

BASE COUNT

ORIGIN

8 others

Query Match		53.8%;	Score 990.8;	DB 9;	Length 1025;
Best Local Similarity		98.6%;	Pred. No. 1.6e-151;		
Matches 1011;		Conservative	7;	Mismatches	5;
				Indels	2;
				Gaps	2;
Qy	777	ACTGGTTCTGGGGTGGACAGCTGGACCTGTCGACGGGACCGCTTGGCGGACAGG	836		
Db	1024	ACTGTTCTGGTGGCGTGGACAGCTGGACCTGTCGACGGGACCGTGGCGGACAGG	965		
Qy	837	CGGCGCTGGACCTGTACCCCTAGCACGCCGGGACGACAGCGGCTTCACCTTCTCCTCCC	896		
Db	964	CGGCGCTGGACCTGTACCCCTAGCACGCCGGGACGACAGC-GCTTACCTTCTCCTCCC	906		
Qy	897	CCAACTTGGCCACATCCCGACGACACGCTGACCGAGATAACGTCTCTCTCCACGC	956		
Db	905	CAAACTTGGCCACATCCCGACGACACGCTGACCGAGATAACGTCTCTCTCCACGC	846		
Qy	957	ACCGGCCAACTCCTTCTACTACCCCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA	1016		
Db	845	ACCGGCCAACTCCTTCTACTACCCCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA	786		
Qy	1017	CACGTGTGCGGCTGCGACAGAGCCCGAGGGCTTTCATCCCTCCCGCCCAAGTCTGCCCCA	1076		
Db	785	CACGTGTGCGGCTGCGACAGAGCCCGAGGGCTTTCATCCCTCCCGCCCAAGTCTGCCCCA	726		
Qy	1077	GCAGGCAATAGATTGTAGACAGCGCCCTCAGTTCCAGAAACGCGCTGGACTGCGAGG	1136		
Db	725	GCAGGCAATAGATTGTAGACAGCGCCCTCAGTTCCAGAAACGCGCTGGACTGCGAGG	666		
Qy	1137	TCCTCCCTGTGCTGCTCTGGGACTGTGCGAGGCGCACTGTGGAGGCTCGGGACCAAGA	1196		
- Db	665	TCCTCCCTGTGCTGCTCTGGGACTGTGCGAGGCGCACTGTGGAGGCTCGGGACCAAGA	606		
Qy	1197	GCAGGACTCGCTACGTCGCGGTCCAGCCCGCCCAACACGAGCGCCCTGCCCGAGCTCG	1256		
Db	605	GCAGGACTCGCTACGTCGCGGTCCAGCCCGCCCAACACGAGCGCCCTGCCCGAGCTCG	546		
Qy	1257	AAGAAGGCTGAGTGGTCCCTGTAACTCGTCTTAAGACCAAGAGCCCGCCAGCCCTG	1316		
Db	545	AAGAAGGCTGAGTGGTCCCTGTAACTCGTCTTAAGACCAAGAGCCCGCCAGCCCTG	486		
Qy	1317	GGCCCCCGGAGCCATGGGGTGTGCGGGGCTCTGTGAGGCTCATGTGAGGGGGCC	1376		
Db	485	GGCCCCCGGAGCCATGGGGTGTGCGGGGCTCTGTGAGGCTCATGTGAGGGGGCC	426		
Qy	1377	GAGGGCACAGGGGTTTCGCGCTGCTCCTGACCGGCTGAGGCGCGCCGACCATCTCTG	1436		
Db	425	GAGGSAACAGGGGTTKCGCGCTGCTCCTGACCGGCTGAGGCGCGCCGACCATCTCTG	366		
Qy	1437	CACGAAAGGCGCTCTGGTGGCGGCGACGGGCATTGGGAACAGCCCTCTCTTCCCAA	1496		
Db	365	CACGAAAGGCGCTCTGGTGGCGGSAAGGGCATTTGGGAACAGCCCTCTCTTCCCAA	306		
Qy	1497	CCTTGCTTCTTAGGGGCCCCGCTGTCCTGCTCAGCCCTCTCTCTCTCTCTCTCTCT	1556		
Db	305	CCTTGCTTCTTAGGGGCCCCGCTGTCCTGCTCTCAGCCCTCTCTCTCTCTCTCTCT	246		
Qy	1557	AAGTCATCCCCAAGGCTCCAGCTACTTAATTAATGCTCTTAATTAATGCTCTCTCT	1616		
Db	245	AAGTCATCCCCAAGGCTCCAGCTACTTAATTAATGCTCTTAATTAATGCTCTCTCT	186		
Qy	1617	CAGGAGATTGCTTCATGTCGAGGGGCTGTGCTCCAGCTGTGTCAGATACCTCAGA	1676		
Db	185	CAGGAGATTGCTTCATGTCGAGGGGCTGTGCTCCAGCTGTGTCAGATACCTCAGA	126		
Qy	1677	CCTGTGCTCTAGGCTGTGCTGAGCCCACTCTCCGAGGGCGCATCCAAAGGGGGCCAC	1736		
Db	125	CCTGTGCTCTAGGCTGTGCTGAGCCCACTCTCTCCGAGGG-GCAGGCCAAGCGGGCGAC	67		
Qy	1737	TTGAGAAGTGAATAAATGGGGGCTTTCGGAAGCGTCAGTGTTCATGTTATGGATCTC	1796		
Db	66	TTGAGAAGTGAATAAATGGGGGCTTTCGGAAGCGTCAGTGTTCATGTTATGGATCTC	7		
Qy	1797	TCCTGC	1801		

Db 6 TCTGC 2

RESULT 2
AL554923

LOCUS
DEFINITION

AL554923 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI087YL11 5

prime, mRNA sequence.

ACCESSION
AL554923

VERSION
AL554923.1

KEYWORDS
EST.

SOURCE
human.

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 1009)

AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE
Full-length cDNA libraries and normalization

JOURNAL
Unpublished (2001)

COMMENT
Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .1009

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DI087YL11"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 155 a 360 c 326 g 157 t 11 others

ORIGIN

Query Match 53.3%; Score 981.2; DB 9; Length 1009;

Best Local Similarity 98.2%; Pred. No. 5.6e-150;

Matches 991; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

Qy 162 CGTGCAGCATCGAAGACAGGAGAACTGGAGCCTCATTTGCCGCGCGCGCGCGCGCT 221

Db 1 CGTGCAGCATCGAAGACAGGAGAACTGGAGCCTCATTTGCCGCGCGCGCGCGCT 60

Qy 222 CGGGCTTAATAGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGCGCGCTCCC 281

Db 61 CGGGCTTAATAGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGCGCGCTCCC 120

Qy 282 GCTGCTCTCTCCGGGTGATGGAACCCAGCCGCGCGCGCGCTTGGGCAAGGCCCTCT 341

Db 121 GCTGCTCTCTCCGGGTGATGGAACCCAGCCGCGCGCGCTTGGGCAAGGCCCTCT 180

Qy 342 GCCT 401

Db 181 GCCT 240

Qy 402 GTTCCGCGCAGAGCCCGGCGCAATACAGCATCACCITTCAGGGCAAGTGGAGCAGACGG 461

Db 241 GTWCCGCCAGAGCCCGGCGCAATACAGCATCACCITTCAGGGCAAGTGGAGCAGACGG 300

Qy 462 CTTTCCCAAGCAGTAGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 521

Db 301 CTTTCCCAAGCAGTAGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360

Qy 522 CCGCCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTACGTACAGGGGCTGC 581


```

Db 841 CGCGCTGGACTGCGAGGTCTCCCTGTGCTGCTGGGGAGTGTGGGCGG-CACTGTG 899
Qy 1179 GGAGGCTGGGACCAAGAGCAGGACTCGCTACGTCGGGTCCAGCCGCCCAACACGGGA 1238
Db 900 GGAGGCTGGGACCAAGAGCAGGACTCGCTACGTCGGGTCCAGCCGCCCAACACGGGA 959
Qy 1239 GCCC 1242
Db 960 GCCC 963

RESULT 4
AL564981/C
LOCUS AL564981 LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN002YN20 3
DEFINITION prime, mRNA sequence.
ACCESSION AL564981
VERSION AL564981.1 GI:12915901
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 989)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..989
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN002YN20"
/clone_lib="LTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was constructed by Life Technologies.
Contact : Feng Liang Life Technologies, a division of
Invitrogen 9800 Medical Center Drive Rockville, Maryland
20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 196 a 278 c 338 g 154 t 23 others
ORIGIN

Query Match 50.2%; Score 924; DB 9; Length 989;
Best Local Similarity 95.6%; Pred. No. 1.le-140;
Matches 943; Conservative 23; Mismatches 18; Indels 2; Gaps 2;

Qy 750 TTGTGTGCGCATCGTCCACAGCCCGGCTGGTTCGTGGCGTGACAGCGCTGGACCTGT 809
Db 986 TTTTGTGCGCWTGTCGCCAGCCCAACTGGTTCGTGGCGTGACAGCGCTGGACCTGT 927
Qy 810 GCGACGGGACCGTTGGCGGGAACAGCGCGCTGGACCTGTACCCCTAGACGCCGGGA 869
Db 926 GCRACGGGACCGTTGGCGGGAACAGCGCGCTGGACCTGTACCCCTAGACGCCGGGA 867
Qy 870 CGGACAGCGGCTTCCACCTTCTCCT-CCCCAACCTGCCACCATCCGCGAGGACACGGTG 928
Db 866 CGRACAGGGCTTACCTTCTCCTCCCCAACCTGCCACCATCCGCGAGGACACGGTG 807
Qy 929 ACCGAGATAAGTCTCCTCTCCAGCCACCGGCCAACTCCTTCTACTACCGCGGCTG 988
Db 806 ACCGAGATAAGTCTCCTCTCCTCCAGCCACCGGCCAACTCCTTCTACTACCGCGGCTG 747
Qy 989 AAGGCCCTGCCTCCCATCGCCAGGGTGACACTGGTGGCGGTGCGACAGAGCCCGAGGGCC 1048

```

```

Db 746 AAGGCCCTGCCTCCCATCGCCAGGGTGACACTGCTGCGGCTGCGACAGAGCCCGAGGCC 687
Qy 1049 TTCTATCCCTCCCGGCCAGTCTCTGCCAGCAGGAGCAATGAGATGTTAGACAGCGCCTCA 1108
Db 686 TTCTATCCCTCCCGGCCAGTCTCTGCCAGCAGGAGCAATGAGATGTTAGACAGCGCCTCA 627
Qy 1109 GTTCCAGAAACGCCGCTGGAAGTCTCCCTGTGCTGCTGCTGGGGAGTGTGCGGA 1168
Db 626 GTTCCAGAAACGCCGCTGGAAGTCTCCCTGTGCTGCTGCTGGGGAGTGTGCGGA 567
Qy 1169 GGCACACTGTGGAGGCTCGGACCAAGAGCAGGACTCGCTACGTCGGGTCCAGCCCGCC 1228
Db 566 GGCACACTGTGGAGGCTCGGACCAAGAGCAGGACTCGCTACGTCGGGTCCAGCCCGCC 507
Qy 1229 AACACGCGGACCCCTGCCCGAGCTCGAAGAGGCTGAGTGGTCCCTGATAACTGC 1288
Db 506 AACACGCGGACCCCTGCCCGAGCTCGAAGAGGCTGAGTGGTCCCTGATAACTGC 447
Qy 1289 GTCTAAGACACAGAGCCCGCAGCCCTGGGGCCCGGAGCCATGGGGTGTCTCGGGGCT 1348
Db 446 GTCTAAGACACAGAGCCCGCAGCCCTGGGGCCCGGAGCCATGGGGTGTCTCGGGGCT 387
Qy 1349 CCTGTGAGGCTCATGTGTCAGGCGCGGAGGCGCACAGGGGTTTCGCGCTGCTCTGAC 1408
Db 386 CCTGTGAGGCTCATGTGTCAGGCGCGGAGGCGCACAGGGGTTTCGCGCTGCTCTGAC 327
Qy 1409 CGCGGTGAGCGCGCGGACCATCTCTGCACTGAAGGGCCCTCTGGTGGCGGACAGGGC 1468
Db 326 CGCGGTGAGCGCGCGGACCATCTCTGCACTGAAGGGCCCTCTGGTGGCGGACAGGGC 267
Qy 1469 ATTGGGAAACAGCCTCTCTCTTCCCAACCTTGTCTTAGGGGCCCGCGGTGCCGTCT 1528
Db 266 ATTGGGATACAGCCTCTCTCTTCTCYWCCCTGCTTTTAGGGGCCCGCGGTGCCGTCT 207
Qy 1529 GCTCTCAGCCTCTCTCTCTCTCAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCTAAAT 1588
Db 206 GCTTTWGGCTCTCTCTCTCTCAGGATAAAGTCAATCCCAAGTCTCTCTCTCTCTCTCT 147
Qy 1589 TATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1647
Db 146 TTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 87
Qy 1648 GGTCTCCACGCTGGTGTGAGATACCTCAGACTGTGCTCTAGGCTGTGCTGAGCCACTC 1707
Db 86 AGCTCCACGCTGGTGTGAGATACCTCAGACTGTGCTCTAGGCTGTGCTGAGCCACTC 27
Qy 1708 TCCGAGGGCGCATCTCAAGCGGGGC 1733
Db 26 TCCGAGGGCGCATCTCAAGCGGGGC 1

```

```

RESULT 5
AL550317
LOCUS AL550317 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI039YK16 5
DEFINITION prime, mRNA sequence.
ACCESSION AL550317
VERSION AL550317.1 GI:12887174
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 971)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers

FEATURES

```

```
source
1. .971
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1039YK16"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 147 a 354 c 315 g 152 t 3 others
ORIGIN

Query Match 49.4%; Score 909; DB 9; Length 971;
Best Local Similarity 99.08; Pred. No. 3e-138;
Matches 964; Conservative 2; Mismatches 3; Indels 5; Gaps 5;

Qy 46 GACCCGGGAAAGTGCTGGGACGGCGAGTTGGGAAAGCGGACGCCCGCCCGCCG 105
Db 1 GACCCGGGAAAGTGCTGGGACGGCGAGTTGGGAAAGCGGCA-SSCCCGCCGCGCCG 59

Qy 106 CAGCCCTTCTCCCTTCTCCACGTCCTATCTGCTCTCGCTGGAGGCCAGGCCGTG 165
Db 60 CAGCCCTTCTCCCTTCTCCACGTCCTATCTGCTCTCGCTGGAGGCCAGGCCGTG 119

Qy 166 CAGATCGAAGACAGAGGAACTGGAGGCTCATTTGGCGCGCGCGCGCGCTCGG 225
Db 120 CAGATCGAAGACAGAGGAACTGGAGGCTCATTTGGCGCGCGCGCGCGCTCGG 179

Qy 226 CTTAATAGGAGCTCCGGCTCTGGCTGGGACCCGACCGCTGCCGGCGCGCTCCGCTG 285
Db 180 CTTAATAGGAGCTCCGGCTCTGGCTGGGACCCGACCGCTGCCGGCGCGCTCCGCTG 239

Qy 286 CTCCTGCGGGTGATGGAAGAACCCAGCCGCGCGCGCGCGCTGGCAAGGCCCTCTGGC 345
Db 240 CTCCTGCGGGTGATGGAAGAACCCAGCCGCGCGCGCGCGCTGGCAAGGCCCTCTGGC 299

Qy 346 TCTCTCTCTGGCCACTCTCGGCGCGCGCGCGCGAGCTCTTGGGGAGAGTCCATCTGTC 405
Db 300 TCTCTCTCTGGCCACTCTCGGCGCGCGCGCGAGCTCTTGGGGAGAGTCCATCTGTC 359

Qy 406 CGCCAGAGCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGCGCCTT 465
Db 360 CGCCAGAGCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGCGCCTT 419

Qy 466 CCCAAGCAGTACCCCTGTTCCGCGCGCGCGCGAGTGTCTTCGCTGTGGGGCGCGC 525
Db 420 CCCAAGCAGTACCCCTGTTCCGCGCGCGCGCGAGTGTCTTCGCTGTGGGGCGCGC 479

Qy 526 GCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTACGTAAACGGGTCGCGA 585
Db 480 GCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTAAACGGGTCGCGA 539

Qy 586 CTTTCGAGGCGCGGAGGCTTGGGCGCTGATGAAGGAGATCGAGCGCGGGGGAGGC 645
Db 540 CTTTCGAGGCGCGGAGGCTTGGGCGCTGATGAAGGAGATCGAGCGCGGGGGAGGC 599

Qy 646 GCTGACAGCGTGACAGAGGTGTTTTTCGGCGCGCGCGCTCCCGACGCGGACCGGACAG 705
Db 600 GCTGACAGCGTGACAGAGGTGTTTTTCGGCGCGCGCGCTCCCGACGCGGACCGGACAG 659

Qy 706 GTCGCGGAGCTGGAGGTGACGAGCAGTCTGCTGCTCGCTTGTGTTGGTGCATCGT 765
Db 660 GTCGCGGAG-CTGAGAGGTGACGAGCAGTCTGCTGCTCGCTTGTGTTGGTGCATCGT 718

Qy 766 GCCCCAGCCCGACTGTTGCTGGCGGTGGACAGCCTGGACCTGTGCGAGCGGGGACCGTTG 825
Db 766 GCCCCAGCCCGACTGTTGCTGGCGGTGGACAGCCTGGACCTGTGCGAGCGGGGACCGTTG 825

Db 719 GCCCAGCCCGACTGTTGCTGGCGGTGGACAGCCTGTGCGAGCGGGGA-CGTTG 777
Qy 826 GCGGGAACAGCGCGCTGGACCTGTACCCCTACGACGCGCGGACGACGCGCTTCAC 885
Db 778 GCGGGAACAGCGCGCTGGACCTGTACCCCTACGACGCGCGGACGACGCGCTTCAC 837
Qy 886 CTTCTCTCTCCCGCAACTTTCGCCACCATCCCGCAGAGAC-ACGGTGACCGAGATAACGCTCT 944
Db 838 CTTCTCTCTCCCGCAACTTTCGCCACCATCCCGCAGAACACGGTGACCGAGATAACGCTCT 897
Qy 945 CTTCTCTCCAGCACCAGCGGCAACT-CTTCTTACTACCGCGCGCTGAAGGCCCTCGCTCC 1003
Db 898 CTTCTCTCCAGCACCAGCGGCAACTCTCTTACTACCGCGCGCTGAAGGCCCTCGCTCC 957
Qy 1004 ATCCGCAAGGTGAC 1017
Db 958 ATCCGCAAGGTGAC 971

RESULT 6
AL574979/c
LOCUS AL574979 LTI_NFL006_PL2 1011 bp mRNA linear EST 16-FEB-2001
DEFINITION AL574979 prime, mRNA sequence.
ACCESSION AL574979
VERSION AL574979.1 GI:12935698
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization.
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
location/Qualifiers
1. .1011
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1064YC22"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 189 a 295 c 343 g 174 t 10 others
ORIGIN

Query Match 48.9%; Score 899; DB 9; Length 1011;
Best Local Similarity 96.5%; Pred. No. 1.3e-136;
Matches 951; Conservative 8; Mismatches 21; Indels 5; Gaps 4;

Qy 799 CTGGACCTGTGCGACGGGACCGTGGCGGGAACAGCGCGCTGTACCCCTCA 858
Db 1000 CTGGACCTGTGCGACGGGACCGTGGCGGGAACAGCGCGCTGTACCCCTCA 944

Qy 859 CGACCGCGGACGACGCGGCTTACCTCTCTCCCGCAACTTCGCGCACCATCCCGCA 918
Db 943 CGACCGCGGACGACGCGGCTTACCTCTCTCCCGCAACTTCGCGCACCATCCCGCA 884

Qy 919 GGACACGCTGACCGAGATAAGCTCCTCTCTCCAGCCACCGCGCAACTCCTTCTACTA 978
Db 919 GGACACGCTGACCGAGATAAGCTCCTCTCTCCAGCCACCGCGCAACTCCTTCTACTA 978
```

Db	883	GGACCGGTGACCGAGATAACGCTCTCTCTCCACGACACCCGGCCAACTCCTTCTCTACTA	824
Qy	979	-CCCGCGGCTGAAGGCCCTCGCTCCCATCG-CCAGGGTGACACTGGTGCGGCTCGCAGAC	1036
Db	823	CCCGCGGCTGAAGGCCCTCGCTCCCATCGCCACGGGTGACACTGGTGCGGCTCGCAGAC	764
Qy	1037	AGCCCGAGGGCTTCATCCCTCCCGCCCGAGTCTCGCCACGACGGGACAATGAGATTGTA	1096
Db	763	AGCCCGAGGGCTTCATCCCTCCCGCCCGAGTCTCGCCACGACGGGACAATGAGATTGTA	704
Qy	1097	GACAGGCGCTCAGTTCACAAAACGGCGGTGGACTCGGAGGTCTCCCTGTGGTCTCCTCTGG	1156
Db	703	GACAGGCGCTCAGTTCACAAAACGGCGGTGGACTCGGAGGTCTCCCTGTGGTCTCCTCTGG	644
Qy	1157	GGACTGTGGGAGGCCACTGTGTGGAGGTCTGGGACCAAGACGAGACTCGCTACGTCCGG	1216
Db	643	GGACTGTGGGAGGCCACTGTGTGGAGGTCTGGGACCAAGACGAGACTCGCTACGTCCGG	584
Qy	1217	GTCCAGCCGCCAAACCGGAGGCCCTCTGCCCGAGCTCGAAGAAGAGGCTGAGTGGCTC	1276
Db	583	GTCCAGCCGCCAAACCGGAGGCCCTCTGCCCGAGCTCGAAGAAGAGGCTGAGTGGCTC	524
Qy	1277	CCTGATAACTCGCTCTAAGACCAGAGCCCGCAGCCCTCGGGGCCCCCGGAGCCATGGG	1336
Db	523	CCTGATAACTCGCTCTAAGACCAGAGCCCGCAGCCCTCGGGGCCCCCGGAGCCATGGG	464
Qy	1337	GTGTCCGGGGCTCCTGTGCAGGCTCATGGTGCAGCGGCCGAGGGCACAGGGGGTTTCGC	1396
Db	463	GTGTCCGGGGCTCCTGTGCAGGCTCATAGTCGAGCGGCCGAGGGCACAGGGGGTTTGC	404
Qy	1397	GCTGCTCTGACCGCGGTGAGCGCGCGGACACTCTGTCACATGAAGGGCCCTCTGGTG	1456
Db	403	GCTGCTCTGACCGCGGTGAGCGCGCGGACCACTCTGTCACATGAAGGGCCCTCTGGTG	344
Qy	1457	GCCGGCACGGGCATTTGGGAAACAGCCCTCCTCTCTTCCCAACCTCTTCTTAGGGGCCCC	1516
Db	343	GCCGGSACGGGCATTTGGGAAACAGCCCTCCTCKTTCCTCCCACTCTTCTTAGGGGCCCC	284
Qy	1517	CGTGTCCCGCTCTGCTCAGCCCTCCTCTCGCAGGATAAAGTCATCCCCAAGGCTCCA	1576
Db	283	CGTGTCCCGCTCTGCTCAGCCCTCCTCKCTCGCAGGATAAAGTCATCCCCAAGGCTCCA	224
Qy	1577	GCTACTCTAAATTATGCTCTCTTAAAGTTATTGTGCTCCAGAGAGATTGTCTTCATCG	1636
Db	223	GCTACTCTAAATTATGCTCTCTTAAAGTTATTGTGCTCCAGAGAGATTGTCTTCATCG	164
Qy	1637	TCAGGGGCTCGCTCCACAGTGGTTGCAGATACCTCAGACCTTGGTGTCTAGGCTGTGC	1696
Db	163	TCAGGGGCTCGCTCCACAGTGGTTGCAGATACCTCAGACCTTGGTGTCTAGGCTGTGC	104
Qy	1697	TGAGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGCCACTTTGAGAAGTGAATAAATGGG	1756
Db	103	TGAGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGCCACTTTGAGAAGTGAATAAATGGG	44
Qy	1757	GCGGTTTCGGAAGCGTCACTGTTTC	1781
Db	43	GCGGTTTCGGAAGCGTCACTGTTTC	19

TITLE Full-length cDNA libraries and normalization
UNPUBLISHED (2001)
JOURNAL Contact: Genoscope
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source 1. 981
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI069VF21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 186 a 280 c 331 g 173 t 11 others
ORIGIN

Query Match 48.6%; Score 893.4; DB 9; Length 981;
Best Local Similarity 96.8%; Pred. No. 1e-135;
Matches 943; Conservative 9; Mismatches 17; Indels 5; Gaps 4;


```
QY 1477 ACAGCCTCCTCCTTCCCAACCTTGCTTCTTAGGGGCCCCCGGTGCTCCCGTCTCTCTCAG 1536
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 ACAGCCTCCTCCTTCCCAACCTTGCTTCTTAGGGGCCCCCGGTGCTCCCGTCTCTCAG 265

QY 1537 CTTCTCCTCCTCGAGATAAGTATCATCCCAAGGCTCAGCTACCTCTAAATATGTCTC 1596
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 264 CTTCTCCTCCTCGAGATAAGTATCATCCCAAGGCTCAGCTACCTCTAAATATGTCTC 205

QY 1597 CTTTAAAGTTATTGCTCTCCAGAGATTGCTCTCATCGTCCAGGGGCTGCTCCAC 1656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 CTTTAAAGTTATTGCTCTCCAGAGATTGCTCTCATCGTCCAGGGGCTGCTCCAC 145

QY 1657 GTGGTTGCAGATACCTCAGACCTGGTCTCTAGGCTGTGCTGAGCCACTCTCCCGAGGG 1716
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 144 GTGGTTGCAGATACCTCAGACCTGGTCTCTAGGCTGTGCTGAGCCACTCTCCCGAGGG 85

QY 1717 CGATCCAGCGGGGGCGCACTTGAGAAGTGAATAAATGGGGCGGTTTCGGAAGCGTCAGT 1776
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 CGATCCAGCGGGGGCGCACTTGAGAAGTGAATAAATGGGGCGGTTTCGGAAGCGTCAGT 25

QY 1777 GTTTCATGTTATG 1790
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 GTTTCATGTTCTG 11

RESULT 8
AL577524/C 950 bp mRNA linear EST 16-FEB-2001
LOCUS AL577524 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI087YL1 3
DEFINITION prime, mRNA sequence.
ACCESSION AL577524
VERSION AL577524.1 GI:12940739
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 950)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191*91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI087YL1"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 169 a 267 c 327 g 160 t 27 others
ORIGIN
Query Match 48.5%; Score 891.6; DB 9; Length 950;
Best Local Similarity 96.1%; Pred. No. 2e-135;
Matches 914; Conservative 8; Mismatches 27; Indels 2; Gaps 2;

QY 852 ACCCTACGACGGCGGACGACAGCGGTTACCTTCTCTCCCCCACTTCGCCACCA 911
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 950 ACCCTACGACGGCGGACGACAGC-GSTTACCTTCTCTCCCCCACTTCGCCACCA 892
```

```
QY 912 TCCCGCAGGACACGGGTGACCGAGATAACGTTCCTCTCTCCAGCCACCCGGCCAACTCCT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 891 TCCCGCAGRAMAGGTTGACCGAGATAACGTTCCTCTCTCCAGCCACCCGGCCAACTCCT 832

QY 972 TCTACTACCCCGGGCTGAAGGGCCTGCCCTCCCATGCCAGGGTGACACTGGTGCGGCTGC 1031
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 831 TCTACAAACCGCGGCTGAAGGGCCTGCCCTCCCATGCCAGGGTDAAACTGGTGCGGCTGC 772

QY 1032 GACAGAGCCCCAGGGGCTTCATCCCTCCCG-CCCCAGTCTCTGCCAGCAGGGACAATGAG 1090
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 771 GACAGAGCCCCAGGGGCTTCATCCCTCCCGCCCCAGTCTCTGCCAGGACAGACAATGAG 712

QY 1091 ATTGTAGACAGCGCCTCAGTTCCAGAAAGCCGCTGGACTTGCAGAGTCTCTCTGTGGTGC 1150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 711 ATTGTAGACAGCGCCTCAGTTCCAGAAAGCCGCTGSACTTGCAGAGTCTCTCTGTGGTGC 652

QY 1151 TCTTGGGGACTGTGCGGAGGCCACTGTGGAGGCTCGGACCAAGAGAGAGACTTCGCTAC 1210
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 651 TCTTGGGGACTGTGCGGAGGCCACTGTGGAGGCTCGGAGCCNAGGCCAGAACTCGCTAC 592

QY 1211 GTCCGGGTCCAGCCGCCAACACGGAGGCCCTGCCGAGCTCGAAGAGAGGCTGAG 1270
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 591 GTCCGGGTCCAGCCGCCNACAAACGGAGGCCCTGCCGAGCTCGAAGNAGAGGCTGAG 532

QY 1271 TGGTCCCTGATAACTGCGTCTTAAGACACAGAGCCCGCAGCCCTGGGGCCCCCGGAGC 1330
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 531 TGGTCCCTGATNACTTGGTCTTAAGACACAGAGCCCGCAGCCCTGGGGCCCCCGGAGC 472

QY 1331 CATGGGTGTGCGGGGCTCCTGTGCAGGCTCATGTCTCAGGGCGCCGAGGGCACAGGGG 1390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 471 CATGGGTGTGCGGGGCTCCTGTGNGCAGGCTCANACNGCAGGGCGCCGAGGNACAGGGG 412

QY 1391 TTTCGGCTGTCTCTGACCGCGGTGAGGGCGCCGACCATCTCTGCAGTGAAGGGCCCT 1450
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 411 TTNTGGCGNGCTCTTGACCGCGGTGAGGGCGCCGACCATNTCTGCAGTGAAGGGCCCT 352

QY 1451 CTGCTGGCGGGCACGGGCTTGGGAAACAGCCCTCTCTTTCCCAACCTTGTCTTAGG 1510
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 351 CTGCTGGCGGGCACGGGCTTGGGANACAGCCCTCTCTTTCCCAACCTTGTCTTAGG 292

QY 1511 GGCCCCGTGTCCCGTCTGTCTCAGCCCTCTCTCTCGAGGATAAAGTCATCCCCAAG 1570
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 291 GGCCCCGTGTCCCGTCTGTCTCAGCCCTCTCTCGAGGATAAAGTCATCCCCAAG 232

QY 1571 GCTCCAGCTACTCTAAATATGTCTCTTATAAGTTATTGCTGCTCAGAGATTGCTCT 1630
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 GCTCCAGCNACTCTAAATATGTCTCTTATAAGTTATTGCTGCTCAGAGATTGCTCT 172

QY 1631 TCATCTCCAGGGGCTGCTCCAGCTGTTGTCAGATACCTCAGACCTCGTCTCTAGG 1690
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 TCATCTCCAGGGGCTGCTCCAGCTGTTGTCAGATACCTCAGACCTCGTCTCTAGG 112

QY 1691 CTGCTGAGCCCACTCTCCCGAGGGCGCATCCAAGGGGGCCACTTGAAGAAGTGAATA 1750
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 CTGCTGAGCCCACTCTCCCGAGGGCGCATCCAAGGGGGCCACTTGAAGAAGTGAATA 52

QY 1751 AATGGGCGGTTTCGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGC 1801
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 AATGGGCGGTTTCGAAGCGTCAGTGTTCATGTTATGGATCTCTCTCSC 1

RESULT 9
AL532820 935 bp mRNA linear EST 13-FEB-2001
LOCUS AL532820 LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN002YN20 5
DEFINITION prime, mRNA sequence.
ACCESSION AL532820
VERSION AL532820.1 GI:12796313
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 935)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source
1. .935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN002YN20"
/clone_lib="LTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/note="Vector: pcMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcMVSPORT 6 vector. Library was constructed by Life Technologies, Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 138 a 328 c 312 g 145 t 12 others
ORIGIN

Query Match 47.4%; Score 871.4; DB 9; Length 935;
Best Local Similarity 98.0%; Pred. No. 3.8e-132;
Matches 910; Conservative 12; Mismatches 3; Indels 4; Gaps 4;

QY 23 CTGCCAGGGGAAGAGGTGATCGACCCGGGAAGTCGCTGGCAGGCGGAGTTGGGAA 82
Db 4 CTGCCAGGGGAAGAGGTGATCGACCCGGGAAGTCGCTGGCAGGCGGAGTTGGGAA 63

QY 83 AGCGGACGCCCGCGCCGCCAGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 142
Db 64 AGCGGCA-SSCG 122

QY 143 CTCTCGTGTGAGCGCAGCGCGTGCAGATCGAAGACAGGAGGAACATGGAGCCCTCATTTGGC 202
Db 123 CTCTCGTGTGAGCGCAGCGCGTGCAGATCGAAGACAGGAGGAACATGGAGCCCTCATTTGGC 182

QY 203 CGGCCCGGGCGCGCGCTCGGGCTTAAATAGAGAGCTCGCGGGCTCTGGCTGGGACCCGAC 262
Db 183 CGGCCCGGGCGCGCGCTCGGGCTTAAATAGAGAGCTCGCGGGCTCTGGCTGGGACCCGAC 242

QY 263 CGCTGCGCGCGCGCGCT 322
Db 243 CGCTGCGCGCGCGCGCT 302

QY 323 GCCTTGGGCAAGGCGCTCTCGCT 382
Db 303 GCCTTGGGCAAGGCGCTCTCGCT 362

QY 383 CTTTGGGGAGAGTCCATCTCTTCCGCCAGAGCCCGGCCCAATACAGCATCACCCTTCACG 442
Db 363 CTTTGGGGAGAGTCCATCTCTTCCGCCAGAGCCCGGCCCAATACAGCATCACCCTTCACG 422

QY 443 GGCAAGTGGAGCCAGAGCGGCTTCCCAAGCAGTACCCCTGTTCCGCCCGCCCTCTGCGCAG 502
Db 423 GGCAAGTGGAGCCAGAGCGGCTTCCCAAGCAGTACCCCTGTTCCGCCCGCCCTCTGCGCAG 482

QY 503 TGGTCTTCGTGTGGGGCGCGCATAGTCCGACTACAGCATGTGGAGGAAGAACCCAG 562
Db 483 TGGTCTTCGTGTGGGGCGCGCATAGTCCGACTACAGCATGTGGAGGAAGAACCCAG 542

QY 563 TAGCTCAGTAACGGGTGCGGACTTTGCGGAGCGCGCGAGGCCCTGGGGCGCTGATGAAG 622
Db 543 TAGCTCAGTTACGGGTGCGGACTTTGCGGGA-CGCGGGGAGAGCCCTGGGGCGCTGATGAAG 601

QY 623 GAGATCAGAGCGCGGGGAGGCGCTGCAGAGCGTGCACAGAGGTGTTTTCGGCGCCCGCC 682
Db 602 GAGATCAGAGCGCGGGGAGGCGCTGCAGAGCGTGCACAGAGGTGTTTTCGGCGCCCGC-C 660

QY 683 GTCCCCAGGCGGCACCGGAGAGAGCTCGGCGGAGCTGAGGTGAGCGCAGGCACTCGCTG 742
Db 661 GTCCCCAGGCGGCACCGGAGAGAGCTCGGCGGAGTGGGAGGTGAGCGCAGGCACTCGCTG 720

QY 743 GTCTCCTTTGTGTGCGCATCGTGGCCAGCCGCGCTGCTGCTGGCGTGGGACAGCCTG 802
Db 721 GTCTCCTTTGTGTGCGCATCGTGGCCAGCCGCGCTGCTGCTGGCGTGGGACAGCCTG 780

QY 803 GACCTGTGGGACGCGGACCGCTTGGCGGGAACAGCGCGCTGACCTGTACCCCTACGAC 862
Db 781 GACCTGTGGGACGCGGACCGCTTGGCGGGAACAGCGCGCTGACCTGTACCCCTWCGMC 840

QY 863 GCCGGGACGAGCAGCGGCTTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 922
Db 841 GCCGGGACGAGCAGCGGCTTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 899

QY 923 ACGGTGACCGAGATAACGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 951
Db 900 ACGGTGACCGAGAAACGCT 928

RESULT 10
AL552365 934 bp mRNA linear EST 16-FEB-2001
LOCUS AL552365 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI069YF21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL552365
VERSION AL552365.1 GI:12891195
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 934)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .934
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI069YF21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pcMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 149 a 327 c 303 g 151 t 4 others
ORIGIN

Query Match 47.3%; Score 870; DB 9; Length 934;
Best Local Similarity 98.7%; Pred. No. 6.4e-132;
Matches 925; Conservative 3; Mismatches 4; Indels 5; Gaps 5;

QY 283 CTGCTCCTGCCGGGTGATGAAACCCAGCCCGCGCGCGCCCTGGGCAAGGCCCTCTG 342
Db 1 CTGCTCCTGCCGGGTGATGAAACCCAGCCCGCGCGCGCCCTGGGCAAGGCCCTCTG 60

QY	343	CGCTCTCCTCTCGCCACTCTCGCGCGCGCGCCAGCCTTTGGGGAGAGTCCATCTG	402
Db	61	CGCTCTCCTCTCGCCACTCTCGCGCGCGCGCCAGCCTTTGGGGAGAGTCCATCTG	120
QY	403	TTCCGGCAGAGCCCGGCCAATATACAGCATACCTTTACAGGGCAAGTGGAGCCAGCGG	462
Db	121	TTCCGGCAGAGCCCGGCCAATATACAGCATACCTTTACAGGGCAAGTGGAGCCAGCGG	180
QY	463	CTTCCCAAGCAGTACCCCTGTTCGCGCCCGCTGCGAGTGTCTTCGCTGCTGGGGC	522
Db	181	CTTCCCAAGCAGTACCCCTGTTCGCGCCCGCTGCGAGTGTCTTCGCTGCTGGGGC	240
QY	523	CGCGCATAGCTCGCACTACAGCATGTGGAGAAAGACCACTACCTCAGTAACCGGCTGCG	582
Db	241	CGCGCATAGCTCGCACTACAGCATGTGGAGAAAGACCACTACCTCAGTAACCGGCTGCG	300
QY	583	CGACTTTGGGAGCGCGGAGCGCTTGGGCGCTGTATGAAGAGATCGAGCGCGGGGGA	642
Db	301	CGACTTTGGGAG-GCGGCGAGCGCTTGGGCGCTGTATGAAGAGATCGAGCGCGGGGA	359
QY	643	GGCGCTGCAGAGCGTGCAGAGGTGTTTTCGCGCGCGCGCTCCAGCGGCACCGGCA	702
Db	360	GGCGCTGCAGAGCGTGCAGCGGNTGTTTTCGCGCGCGCGCTCCAGCGGCACCGGCA	419
QY	703	GAGCTCGGCGGAGTGGAGTGCAGCGCAGCACTCGCTGCTCTGTTGTGTGCGCAT	762
Db	420	GAGCTGCGCGGA-STGGAGTGCAGCGCAGGCACTCGTGGTCTCGTTGTGTGCGCAT	478
QY	763	CGTGCCAGCGCCGACTGTGCTGGGCGTGGACAGCCCTGGACCTGTGCGACGGGACCG	822
Db	479	CGTGCCAGCGCCGACTGTGCTGGGCGTGGACAGCCCTGGACCTGTGCGACGGGACCG	538
QY	823	TTGGCGGGAACAGCGCGCTGACCTGTACCCCTACGACGCGCGGACGACAGCGGCTT	882
Db	539	TTGGCGGGAACAGCGCGCGTGTGACCTGTACCCCTACGACGCGCGGACGACAGCGGCTT	598
QY	883	CACCTTCTCTCCCGCAACTTCCGCCACCATCCCGCAGGACAGCGTGACGAGATAACGTC	942
Db	599	CACCTTCTCTCCCGCAACTTCCGCCACCATCCCGCA-GACAGGTCACGAGATAACGTC	657
QY	943	CTCCTCTCCAGCCACCGCGCCAACTCTCTTACTACCCCGCGCTGAAGGCCCTGCCTCC	1002
Db	658	CTCCTCTCCAGCCACCGCGCCAACTCTCTTACTACCCCGCGCTGAAGGCCCTGCCTCC	717
QY	1003	CATCGCAGGTCACACTGGTGGCTGCGACAGAGCCCGCAGGCGCTTATCCCTCCCGC	1062
Db	718	CATCGCAGGTCACACTGGTGGCTGCGACAGAGCCCGCAGGCGCTTATCCCTCCCGC	777
QY	1063	CCAGTCTCTGCCAGCAGGACAAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCC	1122
Db	778	CCAGTCTCTGCCAGCAGGACAAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCC	837
QY	1123	GCTGGACTGGAGGTCCTCTGTGTGCTGCTGGGACTGTGCGGAGGCCACTGTGGGAG	1182
Db	838	GCTGGACTGGAGGTCCTCTGTGTGCTGCTGGGACTGTGCGGAGG-CACTGTGGGAG	896
QY	1183	GCTCGGACCAAGCAGGACTCGCTACGTCCTGGGTC	1219
Db	897	GCTCGGGA-CAAGACGAGGACTCGCTACGTCGGGGTC	932
RESULT	11		
AL551401			
LOCUS	AL551401	904 bp	mRNA linear EST 16-FEB-2001
DEFINITION	AL551401	lrf1_nfl006_pl2	Homo sapiens cDNA clone CS0D1064YC22 5
ACCESSION	AL551401		prime, mRNA sequence.
VERSION	AL551401.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 904)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..904

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="CS0D1064YC22"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 145 a 319 c 289 g 148 t 3 others
ORIGIN

Query Match 45.6%; Score 838.8; DB 9; Length 904;

Best Local Similarity 98.6%; Pred. No. 7.5e-127;

Matches 895; Conservative 3; Mismatches 5; Indels 5; Gaps 5;

QY 287 TCCTGCGGGTGATGAAACCCACCGCGCGCGCCCTGGGCAAGGCCCTCTCGCT 346

Db 1 TCCTGCGGGTGATGAAACCCACCGCGCGCGCCCTGGGCAAGGCCCTCTCGCT 60

QY 347 CTCTCTCTGCGCACTCTCGCGCGCGCGCGCGCTTTGGGGAGAGTCCATCTCTTC 406

Db 61 CTCTCTCTGCGCACTCTCGCGCGCGCGCGCGCTTTGGGGAGAGTCCATCTCTTC 120

QY 407 GCCAGACCCCGGCCAAATACAGCATCAGCTTCACGGGCAAGTGGAGCCAGCGGCTTC 466

Db 121 GCCAGARCCCGGCCAAATACAGCATCAGCTTCACGGGCAAGTGGAGCCAGCGGCTTC 180

QY 467 CCCAAGCAGTACCCCTGTTCGCGCGCGCGCGCGCTTCGCTGTGCGGGCGCGG 526

Db 181 CCCAAGCAGTACCCCTGTTCGCGCGCGCGCGCGCTTCGCTGTGCGGGCGCGG 240

QY 527 CATAGCTCCGACTACAGCATGTGGAGAAAGACCACTAGTACAGGCTGCGCGAC 586

Db 241 CATAGCTCCGACTACAGCATGTGGAGAAAGACCACTAGTACAGGCTGCGCGAC 300

QY 587 TTTGCGAGCGCGCGGAGCCCTGGGCGCTGATGAAGAGATCGAGCGCGGGGAGGCG 646

Db 301 TTTGCGGA-SGCGCGGAGCCCTGGGCGCTGATGAAGAGATCGAGCGCGGGGAGGCG 359

QY 647 CTGCAGAGCGTGCACAGGTGTTTCGCGCGCGCGCGCTCCAGCGCGCCAGGACG 706

Db 360 CTGCAGAGCGTGCACAGGTGTTTCGCGCGCGCGCGCTCCAGCGCGCCAGGACG 419

QY 707 TCGCGGAGCTGGAGTGCAGCGCAGGCACTCGCTGCTCTGCTGCTGCTGCTGCTG 766

Db 420 TCGGCGGA-CTGAGGTGCAGCGCAGGCACTCGCTGCTCTGCTGCTGCTGCTGCTG 478

QY 767 CCCAGCCCCGACTGTTCTGCGCGCTGGAGAGCCTTGAGACTGTGCGACGGGACCTTGG 826

Db 479 CCCAGCCCCGACTGTTCTGCGCGCTGGAGAGCCTTGAGACTGTGCGACGGGACCTTGG 538

QY 827 CGGGAACAGCGCGCTGAGACCTGTACCCCTAGCAGCGCGGAGCGAGCGGCTTCAAC 886

Db 539 CGGGAACAGCGCGCTGAGACCTGTACCCCTAGCAGCGCGGAGCGAGCGGCTTCAAC 597

```
QY 887 TTCTCTCTCCCCCAACTTGGCCACCATCCCGCAGGACACAGGTGACCGGTAAGATTAAGTCTCTCC 946
|||||
Db 598 TTCTCTCTCCCCCAACTTGGCCACCATCCCGCAGGACACAGGTGACCGGTAAGATTAAGTCTCTCC 657
|||||
QY 947 TCTCCAGCCACCCGCGCACTCTTCTACTACCGCGGCTGAAGCCCTGCTCTCCATC 1006
|||||
Db 658 TCCTCCAGCCACCCGCGCACTCTTCTACTACCGCGGCTGAAGCCCTGCTCTCCATC 717
|||||
QY 1007 GCAGAGGTGACACT-GGTGGGGTGGGACAGAGCCCGCAGGGCCCTTTCATCCCTCCCGCC 1065
|||||
Db 718 GCAGAGGTGACACTGGGTGGGGTGGGACAGAGCCCGCAGGGCCCTTTCATCCCTCCCGCC 777
|||||
QY 1066 AGTCCTGCCAGCAGGACAAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCCGT 1125
|||||
Db 778 AGTCCTGCCAGCAGGACAAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCCGT 837
|||||
QY 1126 GGACTCGGAGGTCTCCCTGTGTCGTCCTGGGACACTGTCGGAGGCCACTGTGGAGGCT 1185
|||||
Db 838 GGACTCGGAGGTCTCCCTGTGTCGTCCTGGGACACTGTGCGGA-GSACTGTGGAGGCT 896
|||||
QY 1186 CGGGACCA 1193
|||||
Db 897 CGGGACAA 904
|||||

RESULT 12
AL577445/c 854 bp mRNA linear EST 16-FEB-2001
LOCUS AL577445 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI086YG11 3
DEFINITION prime, mRNA sequence.
ACCESSION AL577445
VERSION AL577445.1 GI:12940581
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source Location/Qualifiers
1..854
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI086YG11"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 168 a 248 c 288 g 145 t 5 others
ORIGIN

Query Match 45.4%; Score 836.2; DB 9; Length 854;
Best Local Similarity 98.9%; Pred. No. 2e-126;
Matches 838; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 935 ATAAAGTCTCTCTCCAGCACCAGCGCAACTCTTACTACCGCGGCTGAAGGCC 994
|||||
Db 854 ATAAAGTCTCTCTCCAGCACCAGCGCAACTCTTACTACWACCGCGGCTGAAGGCC 795
|||||
```

```
QY 995 CTGCTCTCCCATCGCCAGGTGACACTGTGTGGCTGCGACAGAGCCCCAGGCGCTTCATC 1054
|||||
Db 794 CTGCTCTCCCATCGCCAGGTGACACTGTGTGGCTGCGACAGAGCCCCAGGCGCTTCATC 735
|||||
QY 1055 CTTCCCGCCCCAGTCTCTGCCAGCAGGGAATGAGATTGTAGACAGCGCTCAGTTCCA 1114
|||||
Db 734 CTTCCCGCCCCAGTCTCTGCCAGCAGGGAATGAGATTGTAGACAGCGCTCAGTTCCA 675
|||||
QY 1115 GAAACCCCTGGACTGCGAGGTCTCCCTGTGTGCTCTCTGGGGACTGTGCGGAGGCCAC 1174
|||||
Db 674 GAAACCCCTGGACTGCGAGGTCTCCCTGTGTGCTCTCTGGGGACTGTGCGGAGGCCAC 615
|||||
QY 1175 TGTGGAGGCTCGGGACCAAGAGCACTGCTAGCTCCGGTCCAGCCCCGCAACAAC 1234
|||||
Db 614 TGTGGAGGCTCGGGACCAAGAGCACTGCTAGCTCCGGTCCAGCCCCGCAACAAC 555
|||||
QY 1235 GGGAGCCCTGCCCGGAGCTCGAAGAGGCTGAGTGGTCCCTGATTAACCTGCTCTAA 1294
|||||
Db 554 GGGAGCCCTGCCCGGAGCTCGAAGAGGCTGAGTGGTCCCTGATTAACCTGCTCTAA 495
|||||
QY 1295 GACCAGAGCCCGCAGCCCTTGGGGCCCCCGGAGCATGGGGTGTGCGGGGCTCTCTGTG 1354
|||||
Db 494 GACCAGAGCCCGCAGCCCTTGGGGCCCCCGGAGCATGGGGTGTGCGGGGCTCTCTGTG 435
|||||
QY 1355 CAGGCTCATGCTGCAAGCGCGGCGAGGCAACAGGGGTTTTCGCGCTGCTCTGACCGCGT 1414
|||||
Db 434 CAGGCTCATGCTGCAAGCGCGGCGAGGCAACAGGGGTTTTCGCGCTGCTCTGACCGCGT 375
|||||
QY 1415 GAGGCGCGCGGACCATCTCTGCACCTGAAGGCCCTCTGGTGGCCGCGCAGGCAATTGGG 1474
|||||
Db 374 GAGGCGCGCGGACCATCTCTGCACCTGAAGGCCCTCTGGTGGCCGCGCAGGCAATTGGG 315
|||||
QY 1475 AAACAGCTCTCTCTTTCACACTTGTCTTAGGGGCCCCCGTCTCCCTGCTCTGCTCTC 1534
|||||
Db 314 AAACAGCTCTCTCTTTCACACTTGTCTTAGGGGCCCCCGTCTCCCTGCTCTGCTCTC 255
|||||
QY 1535 AGCTCTCTCTCTCTGCAAGATAAAGTCATCCCAAGGCTCCAGCTACTCTAAATATGTC 1594
|||||
Db 254 AGCTCTCTCTCTCTGCAAGATAAAGTCATCCCAAGGCTCCAGCTACTCTAAATATGTC 195
|||||
QY 1595 TCCTTATAAGTTATTGCTGCTCCAGAGATTGTCTTCATCTCCAGGGGCTGGCTCCC 1654
|||||
Db 194 TCCTTATAAGTTATTGCTGCTCCAGAGATTGTCTTCATCTCCAGGGGCTGGCTCCC 135
|||||
QY 1655 ACCTGTTGCAGATACCTCAGACCTGCTGTAGCTGTGTGAGCCCACTCTCCCGAG 1714
|||||
Db 134 ACCTGTTGCAGATACCTCAGACCTGCTGTAGCTGTGTGAGCCCACTCTCCCGAG 75
|||||
QY 1715 GCGGCATCCAAAGCGGGGCGCACTTGAGAGTGAATAAATGGGCGGTTTCGGAAGGCTCA 1774
|||||
Db 74 GCGGCATCCAAAGCGGGGCGCACTTGAGAGTGAATAAATGGGCGGTTTCGGAAGGCTCA 15
|||||
QY 1775 GTGTTTC 1781
|||||
Db 14 GTTTC 8
|||||

RESULT 13
BG741018
LOCUS BG741018
DEFINITION BG741018.1 GI:14051671
ACCESSION BG741018
VERSION BG741018
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
```

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-re@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10637 row: j column: 11
High quality sequence stop: 823.

FEATURES

source

Location/Qualifiers
1. .843
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4779754"
/clone_lib="NCL_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCL_CGAP Library."

BASE COUNT 140 a 304 c 238 g 161 t

ORIGIN

Query Match 45.2%; Score 832.2; DB 10; Length 843;
Best Local Similarity 99.6%; Pred. No. 8.8e-126;
Matches 834; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 866 GGGACGACAGCGGCTTACCTTCTCTCCGCCCAACTTCCGCCACCACTCCGCCAGGACACG 925
Db 1 GGGACGACAGCGGCTTACCTTCTCTCCGCCCAACTTCCGCCACCACTCCGCCAGGACACG 60
QY 926 GTGACCCAGATACGTCCTCTCTCCAGCACCACCGCCGCAACTCTCTACTACCCGCGG 985
Db 61 GTGACCCAGATACGTCCTCTCTCCAGCACCACCGCCGCAACTCTCTACTACCCGCGG 120
QY 986 CTGAAGGCCCTCGCTCCCATCGCAGGCTGACACTGTGGGCTGCGACAGAGCCCGCAGG 1045
Db 121 CTGAAGGCCCTCGCTCCCATCGCAGGCTGACACTGTGGGCTGCGACAGAGCCCGCAGG 180
QY 1046 GCCTTCATCCTCCGCCGCCAGTCTCTCCGCCAGGACCAATGAGATTGAGACAGCGCC 1105
Db 181 GCCTTCATCCTCCGCCGCCAGTCTCTCCGCCAGGACCAATGAGATTGAGACAGCGCC 240
QY 1106 TCAGTTCAGAAAGCGCGTGGACTGGAGTCTCCCTGTGGTCTCTGGGACCTGTGC 1165
Db 241 TCAGTTCAGAAAGCGCGTGGACTGGAGTCTCCCTGTGGTCTCTGGGACCTGTGC 300
QY 1166 GGAGGCCACTGTGGAGGCTCGGACCAAGAGGAGTCTGCTACGTCCGGGTCCAGCC 1225
Db 301 GGAGGCCACTGTGGAGGCTCGGACCAAGAGGAGTCTGCTACGTCCGGGTCCAGCC 360
QY 1226 GCCAACACGGAGCCCTCGCCCGAGCTCGAAGAGAGGCTGAGTCGCTCCCTGATATAC 1285
Db 361 GCCAACACGGAGCCCTCGCCCGAGCTCGAAGAGAGGCTGAGTCGCTCCCTGATATAC 420
QY 1286 TGGCTCTAAGACACAGAGCCCGAGCCCTTGGGGCCCCCGGACCCATGGGGTGTGCGGG 1345
Db 421 TGGCTCTAAGACACAGAGCCCGAGCCCTTGGGGCCCCCGGAGCCATGGGGTGTGCGGG 480
QY 1346 GCTCCTGTGAGGCTCATGTGACAGGCGCGGAGGGCACAGGGGGTTTCGGCTGTCTCT 1405
Db 481 GCTCCTGTGAGGCTCATGTGACAGGCGCGGAGGGCACAGGGGGTTTCGGCTGTCTCT 540
QY 1406 GACCGCGGTGAGCGCGCGGACCATCTCTGCACTGAAGGCCCTCTGTGGCGGCGACG 1465
Db 541 GACCGCGGTGAGCGCGCGGACCATCTCTGCACTGAAGGCCCTCTGTGGCGGCGACG 600
QY 1466 GGCATTGGGAAACAGCCTCTCTCTTCCCAACCTTGTCTTAGGGGCCCGCGGTGTCGG 1525
Db 601 GGCATTGGGAAACAGCCTCTCTCTTCCCAACCTTGTCTTAGGGGCCCGCGGTGTCGG 660

QY 1526

TCCTGCTCTCAGCCTCCTCTCTCCAGGATAAAGTATCATCCCAAGGCTCCAGCTACTCTA 1585
|||||
Db 661 TCCTGCTCTCAGCCTCCTCTCTCCAGGATACAGTATCCCAAGGCTCCAGCTACTCTA 720
QY 1586 AATTATGCTCTCTTATAAGTTATTGCTGCCAGGAGATTGCTTCATCGTCCAGGGG 1645
|||||
Db 721 AATTATGCTCTCTTATAAGTTATTGCTGCCAGGAGATTGCTTCATCGTCCAGGGG 780
QY 1646 CTGGCTCCCAAGCGTGTGACAGATACCTCAGACCTGGTGTCTAGGCTGTCTGAGCC 1702
|||||
Db 781 CTGGCTCCCAAGCGTGTGACAGATACCTCAGACCTGGTGTCTAGGCTGTCTGAGCC 837

RESULT 14

AL551786

LOCUS

AL551786 LTI_NFL006_PL2 979 bp mRNA linear EST 16-FEB-2001
prime, mRNA sequence.

ACCESSION

AL551786

VERSION

AL551786.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 979)

AUTHORS

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

1. .979

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DI061YD20"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dt) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 166 a 334 c 307 g 156 t 16 others

ORIGIN

Query Match

Best Local Similarity

Matches

314

Ccggccgcccccccttggcgaagggcctctgcgctctctcttgcgacctctcggcgccgc

Db 1 Ccggccgcccccccttggcgaagggcctctgcgctctctcttgcgacctctcggcgccgc

QY 374 Ggcccagcctcttgggggagagtcctctgttccgccagagcccgccgccaataacagc

Db 61 Ggcccagcctcttgggggagagtcctctgttccgccagagcccgccgccaataacagc

QY 434 Acctttacggggcgaagtggagccagagcgcccttccccaaagcagtagtaccctctgttccgccc

Db 121 Acctttacggggcgaagtggagccagagcgcccttccccaaagcagtagtaccctctgttccgccc

QY 494 Cctgcgcagtggtcttgcgctgtggggcgccgcgcatagctccgactacagcatgtggagg

Db 181 Cctgcgcagtggtcttgcgctgtggggcgccgcgcatagctccgactacagcatgtggagg

45.2%; Score 831.6; DB 9; Length 979;
95.8%; Pred. No. 1.1e-125;
Conservative 11; Mismatches 20; Indels 10; Gaps 10;

QY	554	AAGAACCAGTACGTACAGTAAACGGGCTGCGGACACTTTTCGGAGCGCGGAGCGCCTGGGCG	613
Db	241	AAGAACCAGTACGTACAGTAAACGGGCTGCGGACACTTTTCGGAG - GCGGCGAGGCGCTGGGCG	299
QY	614	CTGATGAAGGAGATCAGAGCGCGGGGAGCGCTGCAGAGCGTGCAGAGGTGTTTTCG	673
Db	300	CTGATGAAGGAGATCAGAGCGCGGGGAGCGCTGCAGAGCGTGCAGAGGTGTTTTCG	359
QY	674	GCGGCCGCGCTCCACAGCGGACCGGGCAGACGCTCGGCGGAGCTGGAGTGCACGCGAGG	733
Db	*360	GCGGCCGCGCTCCACAGCGGACCGGGCAGACGCTCGGCGGAG - STGGAGGTGCAGCGCAGG	418
QY	734	CACTCGCTGGTCTCGTTTGTG - GTGCGCATCGTGCCACGCCGACCTGGTTCGTGGGCGT	792
Db	419	CACTCGCTGGTCTCGTTTGTGCGTCGATCGTGCCACGCCGACCTGGTTCGTGGGCGT	478
QY	793	GGACAGCCTGGACCTGTGCGACGGGAGACCGTTTGGCGGGAAACAGGCGGCGTGGACCTGTA	852
Db	479	GGACAGCCTGGACCTGTGCGACGGGAGACCGTTTGGCGGGAAACAGGCGGCGTGGACCTGTA	538
QY	853	CCCTTACCACGCCGGGACGGACAGCGGCTTCACCTTCCTCCCTCCCTCCCTCCCTCCCT	912
Db	539	CCCTTACCAGCGGGGAGGACGGGCTTCACCTTCCT - CCCCRAATTTGNCACAT	597
QY	913	CCCGCAGGACAGGCTGACCGAGATAACCTCCTCTCTCCACGCCACCGGCGCACTCCTT	972
Db	598	CCGNA - GAAACGGTGACCGAGATAACGTCCTCTCTCCACGCCACCGGCGCACTCCTT	656
QY	973	CTACTACCGCGGCTGAAGGCCCTGCTCCCATCGCCAGGTGACACTGGTTCGCGCTGCG	1032
Db	657	CTACTACCGCGGCTGAAGGCCCTGCTCCCATCGCCAGGTGACACTGGTTCGCGCTGCG	716
QY	1033	ACAGAGCCCCAGGCGCTTCATCCCTCCCGCCCCAGTCTCTGCCACGAGGACAATGAGAT	1092
Db	717	ACAGAGCCCCAGGCGCTTCATCCCTCCCGCCCCAGTCTCTGCCACGAGGACAATGAGAT	776
QY	1093	TG - TAGACAGCGCCTCAGTTCCAGAAACGCGCTGGAATGCGAGGTCTCCCTGTGG - TCG	1150
Db	777	TGCTAGACAGCGCCTCAGTTCCAGAAACGCGCTGSACTGCGAGGTCTCCCTGTGGCTG	836
QY	1151	TCCTGGGACACTGTCGGGAGGCCACTG - TGGGAGGCTCGGGACCAAGACGAGCACTCGCTA	1209
Db	837	TCCTGGGACACTGTCGGGAGGCCACTGTCGAGGCTCGSSACCAGACGAGCCTCCTA	896
QY	1210	CGTCGGGTCCAGCGCCGCAACAGCGGAGCGCTCGCCCGAGCTCGAAGAGAGGCTGA	1269
Db	897	GCGCGGCTCAGCGCGCGCAAAAGGAG - CCCTGCCCGAGCTCGAAGAGAGGCTGA	955
QY	1270	GTGGCTCCCTGATAACTGCGTCTAA	1294
Db	956	GTGGCT - CCTATAAATGCGTCTAA	979
RESULT 15			
LOCUS	AL542564		
DEFINITION	AL542564 LTI-FL002_PL1 Homo sapiens CDNA clone CS0DE012YP18 5 prime	904 bp mRNA linear	EST 16-FEB-2001
ACCESSION	AL542564		
VERSION	AL542564.1		
KEYWORDS	GI:12874732		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		

•

```

|||||
Db 719 TGTGTGCGCATCGTGCCAGCCGACTGGTTCTGGCGTGACAGCCTGGACCTGTG 778
QY 811 CGACGGGGACCGTTGGCGGGACAGCGCGGCTGGACCTGTACCCC-TACGACGCGGGG 869
Db 779 CGACGGGGACGTGGCGGGGACAGCGCGGCTGGACCTGTACCCCTTACGACGCGGGG 838
QY 870 CGGACAGCGGCTTACCTTCTCCTCCCGCAACTTCGCCACCATCCCGCAGGACACGGTGA 929
Db 839 CGGACAGCGGCTTACCTTCTCCTCCCGCAACTTCGCCACCATCCCGCAGGACACGGTGA 898
QY 930 CCGAGA 935
Db 899 CCGAGA 904
|||||
```

Search completed: October 27, 2002, 11:14:29
Job time : 2003 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2002, 06:19:17 ; Search time 3403 Seconds
(without alignments)
11314.965 Million cell updates/sec

Title: US-09-807-200-1
Perfect score: 1840
Sequence: 1 ctcgagcgggtgcgcgag.....ctgtgtctcaaaaaaaaaa 1840

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl : *
1: gb_ba : *
2: gb_hgt : *
3: gb_in : *
4: gb_om : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vi : *
15: em_ba : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_mu : *
20: em_on : *
21: em_or : *
22: em_ov : *
23: em_pat : *
24: em_ph : *
25: em_pl : *
26: em_ro : *
27: em_sts : *
28: em_un : *
29: em_vi : *
30: em_htg_hum : *
31: em_htg_inv : *
32: em_htg_other : *
33: em_htgo_inv : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Match Length	Description

1	1764.4	95.9	1807	9	AB027466	AB027466 Homo sapi
2	1761.2	95.7	1785	6	AX203883	AX203883 Sequence
3	1692.4	92.0	1779	6	AR167936	AR167936 Sequence
4	1599.4	86.9	1634	9	BC002707	BC002707 Homo sapi
5	1100.2	59.8	1105	6	AR035961	AR035961 Sequence
6	988	53.7	996	6	BD006592	BD006592 Integrin
7	729.2	39.5	2820	9	AK026054	AK026054 Homo sapi
8	726.4	39.5	190462	9	AC092535	AC092535 Homo sapi
9	726.4	39.5	197176	2	AC022763	AC022763 Homo sapi
10	724	39.3	4292	9	AK024499	AK024499 Homo sapi
11	685.8	37.3	2325	10	AF155196	AF155196 Rattus no
12	679	36.9	1778	10	BC017616	BC017616 Mus muscu
13	549.8	29.9	1020	6	BD006593	BD006593 Integrin
14	364.2	19.8	2116	5	AB006085	AB006085 Danio rer
15	343.8	18.7	400	6	AR037874	AR037874 Sequence
16	339	18.4	400	6	AR125230	AR125230 Sequence
17	304.8	16.6	506	6	AR035966	AR035966 Sequence
18	263.4	14.3	265	6	AX336981	AX336981 Sequence
19	259.6	14.1	3381	5	AB006084	AB006084 Danio rer
20	204.2	11.1	316	6	AR035967	AR035967 Sequence
21	119.8	11.1	316	6	AR035968	AR035968 Sequence
22	119.8	6.5	2565	3	DMAF000178	AF000178 Drosophil
23	110.6	6.0	3489	3	BFL6096	AJ06096 Branchios
24	106.6	5.8	2466	3	AF135119	AF135119 Drosophil
25	88.8	4.8	1786	3	AY060990	AY060990 Drosophil
26	87.2	4.7	125020	9	AF429315	AF429315 Homo sapi
27	81.6	4.4	2836	3	AY058267	AY058267 Drosophil
28	81.2	4.4	3146	5	AB006086	AB006086 Danio rer
29	79.6	4.3	4061	4	AB051389	AB051389 Bos tauru
30	79.6	4.3	4061	6	E32471	E32471 Novel vascu
31	77.8	4.2	51289	3	AC005717	AC005717 Drosophil
32	77.8	4.2	73721	2	AC020007	AC020007 Drosophil
33	77.8	4.2	163072	3	AC099032	AC099032 Drosophil
34	77.8	4.2	259817	3	AE003804	AE003804 Drosophil
35	74.4	4.0	125020	9	AF429315	AF429315 Homo sapi
36	71.6	3.9	3079	6	AX093366	AX093366 Sequence
37	71.6	3.9	3079	6	E32470	E32470 Novel vascu
38	71.6	3.9	3079	9	AB051390	AB051390 Homo sapi
39	71.6	3.9	3999	6	AX093359	AX093359 Sequence
40	71.6	3.9	3999	9	AB018305	AB018305 Homo sapi
41	71.6	3.9	4029	6	AR007441	AR007441 Sequence
42	71.6	3.9	4029	6	E32472	E32472 Novel vascu
43	71.6	3.9	4029	10	RATFSAA	M88469 Rattus norv
44	71.6	3.9	4035	10	BC020531	BC020531 Mus muscu
45	70.6	3.8	76217	2	AC020385	AC020385 Drosophil

ALIGNMENTS

RESULT 1	AB027466	AB027466	1807 bp	mRNA	linear	PRI 05-NOV-1999
LOCUS	Homo sapiens SPON2 mRNA for spondin 2, complete cds.					
DEFINITION	AB027466					
ACCESSION	AB027466.1	GI:6172220				
VERSION	SPON2; spondin 2.					
KEYWORDS	Homo sapiens cDNA to mRNA.					
SOURCE	Homo sapiens					
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (sites)					
AUTHORS	Manda, R., Kohno, T., Matsuno, Y., Takenoshita, S., Kuwano, H. and Yokota, J.					
TITLE	Identification of genes (SPON2 and c20orf2) differentially expressed between cancerous and noncancerous lung cells by mRNA differential display					
JOURNAL	Genomics 61 (1), 5-14 (1999)					
MEDLINE	99443867					
REFERENCE	2 (bases 1 to 1807)					
AUTHORS	Yokota, J., Kohno, T. and Manda, R.					
TITLE	Direct Submission					
JOURNAL	Submitted (18-MAY-1999) Jun Yokota, National Cancer Center Research					


```
QY 1769 GGGTCAGTGTTCATGATCTCTCGGTTTGAATAAAGACTATCTCTGTGCT 1828
|||||
Db 1745 GCGTCAGTGTTCATGATCTCTCGGTTTGAATAAAGACTATCTCTGTGCT 1804
QY 1829 CA 1830
|||
Db 1805 CA 1806

RESULT 2
AX203883 AX203883 1785 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0144291.
ACCESSION AX203883
VERSION AX203883.1 GI:15393341
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1785)
AUTHORS Harkins,R., Parkes,D., Parry,G., Schneider,D.W. and Steinbrecher,R.
TITLE Dna encoding a novel rgl polypeptide
JOURNAL Patent: WO 0144291-A 1 21-JUN-2001;
SCHERING AKTIENGESSELLSCHAFT (DE)
FEATURES
Location/Qualifiers
source 1..1785
/organism="Homo sapiens"
/db_xref="taxon:9606"
296..1291
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC60159.1"
/db_xref="GI:15393342"
/translation="MENPSPAALKKALLATLGAAGPGLGESICSAGAPAKYS
ITFGKWSQTFPPKQYPLFRPAPQWSLLGAHSDYSNMRKNQYVSNGLRDFABRGE
AWALKEITEAGEALQSHVAHVSAPAPVSGTSGTSAELEVORHSLVSFVVRIVPSPD
WFGVDSLDCGDWRQALDLYPDAGTDSGTFSSPNFATIPQDVTVEITSSPD
SHPANFVYPLKALPIIARVTLVRLSPRAFIPAPVLPDSRDNEIVDSASVPTPL
DCEVLSWSGLCGHGCRGLGTSKTRVRVQPANNGSPCELEEEACVPCNV"
BASE COUNT 300 a 614 c 570 g 301 t
ORIGIN

Query Match 95.7%; Score 1761.2; DB 6; Length 1785;
Best Local Similarity 99.8%; Pred No 6.9e-249;
Matches 1763; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 GGGGTGCGGACGACATGCCAGGGGAAGAGGTGATCCGACCCGGGAAGGTGCGTGGCA 68
|||||
Db 6 GGGGTGCGGACGACATGCCAGGGGAAGAGGTGATCCGACCCGGGAAGGTGCGTGGCA 65
QY 69 GGGCGAGTTGGGAAGCGGAGCCCGCGCGCCCGCGAGCCCGCTTCTCTCTTTCTCC 128
|||||
Db 66 GGGCGAGTTGGGAAGCGGAGCCCGCGCGCCCGCGAGCCCGCTTCTCTCTTTCTCC 125
QY 129 CAGTCTCTATCTCCCTCTCGTGGAGCCAGGCGCTGCGAGCATCGAAGACAGAGGAAT 188
|||||
Db 126 CAGTCTCTATCTCCCTCTCGTGGAGCCAGGCGCTGCGAGCATCGAAGACAGAGGAAT 185
QY 189 GGAGCCTCATTTGGCCGGCCCGCGCGCTCGGGCTTAATAGAGATCCGGGCTCT 248
|||||
Db 186 GGAGCCTCATTTGGCCGGCCCGCGCGCTCGGGCTTAATAGAGATCCGGGCTCT 245
QY 249 GGTGGACCCGACCGTGCAGCGCGCGCTCCCGCTGCTCTCTCTCTCTCTCTCTCTCT 308
|||||
Db 246 GGTGGACCCGACCGTGCAGCGCGCGCTCCCGCTGCTCTCTCTCTCTCTCTCTCTCT 305
QY 309 CCAGCCCGCGCGCGCTGGGAAGGCCCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCT 368
|||||
Db 306 CCAGCCCGCGCGCGCTGGGAAGGCCCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCT 365
QY 369 CCAGCCCGCGCGCTGTGGGGAGAGTCCATCTGTTCGCGCAGAGCCCGCGCAATACA 428
|||||
```

```
Db 366 CGCCCGCGGAGCCCTCTTGGGGAGAGTCCATCTGTTCGCGGAGCCCGCGCAATACA 425
QY 429 GCATCACCTTCACGGGCAAGTGGAGCCAGACGGCCCTTCCCAAGCAGTACCCCTGTTC 488
Db 426 GCATCACCTTCACGGGCAAGTGGAGCCAGACGGCCCTTCCCAAGCAGTACCCCTGTTC 485
QY 489 GCCCCCTCGCAGTGTCTTCGCTGTGGGGCCCGCATAGCTCCGACTACAGCATGT 548
Db 486 GCCCCCTCGCAGTGTCTTCGCTGTGGGGCCCGCATAGCTCCGACTACAGCATGT 545
QY 549 GGAGGAAGAACAGTACGTACAGTAACGGCTGCGGACCTTTGCGGAGCGCGGAGGCGCT 608
Db 546 GGAGGAAGAACAGTACGTACAGTAACGGCTGCGGACCTTTGCGGAGCGCGGAGGCGCT 605
QY 609 GGGCGCTGATGAAGAGATCGAGGCGCGGGGAGCGGCTGCGAGAGCGTGCACAGAGTGT 668
Db 606 GGGCGCTGATGAAGAGATCGAGGCGCGGGGAGCGGCTGCGAGAGCGTGCACAGCGGTGT 665
QY 669 TTTTCGCGCGCCGCTCCCGAGCGGCACCGGGCAGACGTCGCGGAGCTGGAGGTGCAGC 728
Db 666 TTTTCGCGCGCCGCTCCCGAGCGGCACCGGGCAGACGTCGCGGAGCTGGAGGTGCAGC 725
QY 729 GCAGGCACTCGTGTCTCTGTTGTGTGGCATCGTGCAGCCCGGACTGTTGTTGTTGG 788
Db 726 GCAGGCACTCGTGTCTCTGTTGTGTGGCATCGTGCAGCCCGGACTGTTGTTGTTGG 785
QY 789 GCGTGCACACCTTGGACCTGTGCGACGGGACCGTTGGGGGGAACAGCGGCGCTGGACC 848
Db 786 GCGTGCACACCTTGGACCTGTGCGACGGGACCGTTGGGGGGAACAGCGGCGCTGGACC 845
QY 849 TGTACCCCTACGACGCGGAGCGACAGCGGCTTACCTTCTCTCCCGCACTTCGCCCA 908
Db 846 TGTACCCCTACGACGCGGAGCGACAGCGGCTTACCTTCTCTCCCGCACTTCGCCCA 905
QY 909 CCATCCCGCAGGACACGGTGAACGAGATAACGTCCTCTCTCCAGCCACCCCGCCCAACT 968
Db 906 CCATCCCGCAGGACACGGTGAACGAGATAACGTCCTCTCTCCAGCCACCCCGCCCAACT 965
QY 969 CCTTCTACTACCGCGGCTGAAGGCCCTCGCTCCCATCGCCAGGCTGACACTGTTGGCG 1028
Db 966 CCTTCTACTACCGCACCGAGGCCCTCGCTCCCATCGCCAGGCTGACACTGTTGGCGC 1025
QY 1029 TGGCAGACAGCCCGCAGGCGCTTCACTCCCTCCCGCCCGAGTCTCGCCAGCAGGACAATG 1088
Db 1026 TGGCAGACAGCCCGCAGGCGCTTCACTCCCTCCCGCCCGAGTCTCGCCAGCAGGACAATG 1085
QY 1089 AGATTGTAGACAGCGCTCAGTTCCAGAAACGCGCTGGAGTTCGAGGTTCTCCCTGTGGT 1148
Db 1086 AGATTGTAGACAGCGCTCAGTTCCAGAAACGCGCTGGAGTTCGAGGTTCTCCCTGTGGT 1145
QY 1149 CGTCTCTGGGAGTGTGGGAGGCCACTGTGGAGGCTCGGGACCAAGAGCAGGACTCGCT 1208
Db 1146 CGTCTCTGGGAGTGTGGGAGGCCACTGTGGAGGCTCGGGACCAAGAGCAGGACTCGCT 1205
QY 1209 ACCTCCGGGTCCAGCGCCGCAACAGCGGAGCCCTGCCCCGAGCTCGAAGAAGAGGCTG 1268
Db 1206 ACCTCCGGGTCCAGCGCCGCAACAGCGGAGCCCTGCCCCGAGCTCGAAGAAGAGGCTG 1265
QY 1269 AGTTCGCTTCCCTGATAACTGCGTCTAAGACACAGAGCCCGCCAGCCCTTGGGGCCCCCGGGA 1328
Db 1266 AGTTCGCTTCCCTGATAACTGCGTCTAAGACACAGAGCCCGCCAGCCCTTGGGGCCCCCGGGA 1325
QY 1329 GCATCGGGTGTGCGGGGCTCTGTGAGGCTCATGCTGAGGGGCGCGGAGGACAGAGG 1388
Db 1326 GCATCGGGTGTGCGGGGCTCTGTGAGGCTCATGCTGAGGGGCGCGGAGGACAGAGG 1385
QY 1389 GGTTCGCGCTGCTCTGACCGGCTGAGGCCGCGCGGACCATCTCTGCACTGAAGGGCC 1448
Db 1386 GGTTCGCGCTGCTCTGACCGGCTGAGGCCGCGCGGACCATCTCTGCACTGAAGGGCC 1445
QY 1449 CTCTGTGGCGGCGACGGGCTTGGGAACAGAGCTCTCTCTTTCCCAACCTTGTCTTTA 1508
|||||
```



```

QY 1563 TCCCAAGGCTCAGCTACTCTAAATATGCTCTCTTATAAGTTATGCTGCTCCAGGAG 1622
D 1497 TCCCAAGGCTCAGCTACTCTAAATATGCTCTCTTATAAGTTATGCTGCTCCAGGAG 1556
QY 1623 ATTGCTCTTATCGTCCAGGGCCCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGT 1682
D 1557 ATTGCTCTTATCGTCCAGGGCCCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGT 1616
QY 1683 GCTCTAGGCTGTGCTAGCCCACTCTCCGAGGGCGCATCCAGGGGGCCACTTGAGA 1742
D 1617 GCTCTAGGCTGTGCTAGCCCACTCTCCGAGGGCGCATCCAGGGGGCCACTTGAGA 1676
QY 1743 ACTGAATAAATCGGGCGGTTTCGGAAGCGTCACTGTTT--CCATGTATGATCTCTCTG 1800
D 1677 ACTGAATAAATCGGGCGGTTTCGGAAGCGTCACTGTTTTCGTCNCGCCCGGATCTCTG 1736
QY 1801 CTTTGAATAAAGACTATCTCTGTTGCTCAAAAAA 1840
D 1737 CGTTTGAATAAAGACTATCTCTGTTGCTCAAAAAA 1776

RESULT 4
BC002707
LOCUS
DEFINITION Homo sapiens, spondin 2, extracellular matrix protein, clone
ACCESSION BC002707
VERSION BC002707.1
KEYWORDS GI:12803740
SOURCE MGC.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg,R.
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing Center (NISC), National Institutes of Health Intramural
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
Lim,M., Maduro,Q.L., Masello,C., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
Zhang,B.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 12 Row: a Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6912681.
Location/Qualifiers
1. .1634
/organism="Homo sapiens"
/db.xref="LocusID:10417"
/db.xref="taxon:9606"
/clone="MGC:3472 IMAGE:3630467"
/tissue_type="Placenta, chorioncarcinoma"
/clone_lib="NIH_MGC_21"
FEATURES
source

```

```

/lab_host="DHI01B-R"
/note="Vector: pOTB7"
69..1064
/codon_start=1
/product="spondin 2, extracellular matrix protein"
/protein_id="AAH02707.1"
/db_xref="GI:12803741"
/translation="MENPSPAAALGKALCALLATLGAAGOLPGGESIC SARAPAKYS
ITFTGKNSOTAPFKQYPLFRPPAOWSHLLGAHSDSYSMWRKNQYVNSGLRDPFAERGE
AWALMKEIEAAGEALQSVHEFSPAVPSGTGTSAELEVRHRLSLVSVFVRIVSPD
WFGVYSLDLCDDRWREQAALDLYPDAGTDSGTFSPNFATIPQDTVEITVSSSP
SHPNSEFYPRLKLPPTIARLQSPRAFIPAPVLPFSRDNIEIVDSAVPTPL
DCEVSLSSWGKLCGHCRLGTGKSTRVVRVPANNGSPCPLEBEAEACVPCV"
BASE COUNT 293 a 550 c 501 g 290 t
ORIGIN
Query Match 86.9%; Score 1599.4; DB 9; Length 1634;
Best Local Similarity 99.9%; Pred. No. 3.5e-225;
Matches 1600; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 240 CCGGGCTCTGGCTGGGACCCGACCGCTCCGGCGCGCTCCGGCTGCTCCTCCGGGTGA 299
D 10 CCGGGCTCTGGCTGGGACCCGACCGCTCCGGCGCGCTCCGGCTGCTCCTCCGGGTGA 69
QY 300 TGGAAAAACCCACCGCGCGCGCTGGCAAGGCCCTCTGGCTCTCTCTCCGGGTGA 359
D 70 TGGAAAAACCCACCGCGCGCGCTGGCAAGGCCCTCTGGCTCTCTCTCCGGGTGA 129
QY 360 CTCTCGGGCGCGCGCGCGCTCTTGGGGAGAGTCCATCTGTTCGGCAGAGCCCCGG 419
D 130 CTCTCGGGCGCGCGCGCGCTCTTGGGGAGAGTCCATCTGTTCGGCAGAGCCCCGG 189
QY 420 CCAATACAGCATCACCTTCACGGGCAAGTGGAGCAGAGCGCTTCCCAAGCATACC 479
D 190 CCAATACAGCATCACCTTCACGGGCAAGTGGAGCAGAGCGCTTCCCAAGCATACC 249
QY 480 CCTGTTCGGCGCGCGCTGGCAGTGTCTTCTGCTGGGGCGCGCGCTAGTCTCGACT 539
D 250 CCTGTTCGGCGCGCGCTGGCAGTGTCTTCTGCTGGGGCGCGCGCTAGTCTCGACT 309
QY 540 ACAGCATGTGGAGGAAGAACACAGTACGTCACTAACGGGTTCGGCGAGCGCG 599
D 310 ACAGCATGTGGAGGAAGAACACAGTACGTCACTAACGGGTTCGGCGAGCGCG 369
QY 600 GCGAGCGCTGGCGCTGATGAAGAGATCAGGCGCGGGGAGGCGCTGCAGAGCTGC 659
D 370 GCGAGCGCTGGCGCTGATGAAGAGATCAGGCGCGGGGAGGCGCTGCAGAGCTGC 429
QY 660 ACGAGTGTTCGGCGCGCGCTCCCGAGCGCACCGGGGAGAGCTCGGGAGCTGG 719
D 430 ACGAGTGTTCGGCGCGCGCTCCCGAGCGCACCGGGGAGAGCTCGGGAGCTGG 489
QY 720 AGGTGAGCGCAGGCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
D 490 AGGTGAGCGCAGGCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
QY 780 GCTTCGTGGCGCTGGACAGCGCTGGACCTGTGCGACGGGAGCGCTTGGCGGAGACGCG 839
D 550 GCTTCGTGGCGCTGGACAGCGCTGGACCTGTGCGACGGGAGCGCTTGGCGGAGACGCG 609
QY 840 CGCTGACCTGTACCTACCTACGACCGCGGACGACAGCGGCTTCACCTTCTCTCTCCCA 899
D 610 CGCTGACCTGTACCTACCTACGACCGCGGAGCGGACGACAGCGGCTTCACCTTCTCTCCCA 669
QY 900 ACTTCGCCACCTATCCCGCGCAGGACAGCGTGACCGAGATAACGCTCTCTCTCCAGCCACC 959
D 670 ACTTCGCCACCTATCCCGCGCAGGACAGCGTGACCGAGATAACGCTCTCTCTCCAGCCACC 729
QY 960 CGGCCAACTCTCTTACTACTACCGCGCGCTGAAGGCCCTGCTCCCATCGCCAGGCTGACAC 1019
D 730 CGGCCAACTCTCTTACTACTACCGCGCGCTGAAGGCCCTGCTCCCATCGCCAGGCTGACAC 789
QY 1020 TGGTGGCGCTGCGACAGAGACCGCCAGCGCTTCATCCTCCCGCCCGAGTCTCTCCCGACGA 1079

```

```
Db 790 TCCTCGGCTGCGACAGACCCAGGGCTTCATCCCTCCCGCCAGTCTCCCGCAGCA 849
Qy 1080 GGGACAAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCCGTGGACTCGAGGTTCT 1139
Db 850 GGGACAAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCCGTGGACTCGAGGTTCT 909
Qy 1140 CCCTGTGTGTCCTCGGAGCTGTGGAGGCACTGTGGAGGCTCGGGACCAAGAGCA 1199
Db 910 CCCTGTGTGTCCTCGGAGCTGTGGAGGCACTGTGGAGGCTCGGGACCAAGAGCA 969
Qy 1200 GGACTCGCTACGTCGCGGTTCAGCCGCGCAACAGGGAGCCCTCGCCCGAGTCGGAAG 1259
Db 970 GGACTCGCTACGTCGCGGTTCAGCCGCGCAACAGGGAGCCCTCGCCCGAGTCGGAAG 1029
Qy 1260 AAGAGCTGAGTGCCTGCTGATAACTGCGTCTAAGACAGAGCCCGCAGCCCTGGG 1319
Db 1030 AAGAGCTGAGTGCCTGCTGATAACTGCGTCTAAGACAGAGCCCGCAGCCCTGGG 1089
Qy 1320 CCCCCGGAGCCATGGGTGTCGGGGCTCTGTGTCAGAGCTCATGTGCGAGCGCCGAG 1379
Db 1090 CCCCCGGAGCCATGGGTGTCGGGGCTCTGTGTCAGAGCTCATGTGCGAGCGCCGAG 1149
Qy 1380 GGCACAGGGGTTTCGCGTGTCTGTGACCGCGGTGAGGCGCGCCGAGCATCTCTGCAC 1439
Db 1150 GGCACAGGGGTTTCGCGTGTCTGTGACCGCGGTGAGGCGCGCCGAGCATCTCTGCAC 1209
Qy 1440 TGAAGGGCCCTCTGTGCGCGCACGGGATGGAACAGAGCTCTCTCTTCCCAACCT 1499
Db 1210 TGAAGGGCCCTCTGTGCGCGCACGGGATGGAACAGAGCTCTCTCTTCCCAACCT 1269
Qy 1500 TGCCTCTTAGGGGCCCCCTGTGCGCTCTGCTCTCAGCTCTCTCTCTGCGAGGATAAG 1559
Db 1270 TGCCTCTTAGGGGCCCCCTGTGCGCTCTGCTCTCAGCTCTCTCTCTGCGAGGATAAG 1329
Qy 1560 TCATCCCCAAGGCTCCAGTACTCTAAATATATGTCTCTCTATAAGTATATGTCTCCAG 1619
Db 1330 TCATCCCCAAGGCTCCAGTACTCTAAATATATGTCTCTCTATAAGTATATGTCTCCAG 1389
Qy 1620 GAGATTGCTCTTCATGCTCCAGGGCTGCTGCTCCAGCTCTCTCTCTGCGAGTACCTCAGACCT 1679
Db 1390 GAGATTGCTCTTCATGCTCCAGGGCTGCTGCTCCAGCTCTCTCTCTGCGAGTACCTCAGACCT 1449
Qy 1680 GGTGCTTAGGCTGCTGAGCCACTCTCCCGAGGGGCGCATCCAAAGGGGGCCACTTG 1739
Db 1450 GGTGCTTAGGCTGCTGAGCCACTCTCCCGAGGGGCGCATCCAAAGGGGGCCACTTG 1509
Qy 1740 AGAAGTGAATAAATGGGCGGTTTCGGAAGCGTCAAGTGTTCATGTTATGATCTCTCT 1799
Db 1510 AGAAGTGAATAAATGGGCGGTTTCGGAAGCGTCAAGTGTTCATGTTATGATCTCTCT 1569
Qy 1800 GCGTTTGAATAAAGACTATCTCTGTGCTCAAAAAA 1840
Db 1570 GCGTTTGAATAAAGACTATCTCTGTGCTCAAAAAA 1610
```

```
RESULT 5
AR035961
LOCUS AR035961 1105 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5871969.
ACCESSION AR035961
VERSION AR035961.1 GI:5952629
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 1105)
Hastings, G. and Dillon, P.J.
TITLE Nucleic acids encoding human neuronal attachment factor-1
JOURNAL Patent: US 5871969-A 1 16-FEB-1999;
FEATURES
Location/Qualifiers
1..1105
/organism="unknown"
```

```
BASE COUNT 178 a 393 c 362 g 172 t
ORIGIN
Query Match 59.8%; Score 1100.2; DB 6; Length 1105;
Best Local Similarity 99.7%; Pred. No. 4.9e-152;
Matches 1102; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 281 CCCTGCTCTGTCGGGCTGATGAAACCCAGCCGCGCGCCCTGGCAAGGCCCTC 340
Db 1 CGCTGCTCTGTCGGGCTGATGAAACCCAGCCGCGCGCCCTGGCAAGGCCCTC 60
Qy 341 TCGCTCTCTCTGTCGGCACTCTCGGCGCGCGCGCGCGCTTTGGGGAGAGTCCATC 400
Db 61 TCGCTCTCTCTGTCGGCACTCTCGGCGCGCGCGCGCGCTTTGGGGAGAGTCCATC 120
Qy 401 TGTTCGCGCAGAGCCCGCGCAATACAGCATCACTTCAACGGCAAGTGGAGCCAGAG 460
Db 121 TGTTCGCGCAGAGCCCTGGCCAAATACAGCATCACTTCAACGGCAAGTGGAGCCAGAG 180
Qy 461 GCCTTCCCAAGCAGTACCCCTGTTCCGCGCCCTGCGAGTGGTCTTCGCTGTGGG 520
Db 181 GCCTTCCCAAGCAGTACCCCTGTTCCGCGCCCTGCGAGTGGTCTTCGCTGTGGG 240
Qy 521 GCGCGCATAGTCTCGAGTACAGCATGTGAGGAAGAACAGTACGTCACTAACGGGCTG 580
Db 241 GCGCGCATAGTCTCGAGTACAGCATGTGAGGAAGAACAGTACGTCACTAACGGGCTG 300
Qy 581 GCGCATTTGCGGAGCGCGCGAGGCTGGGCGCTGATGAAGAGATCGAGCGCGGGG 640
Db 301 GCGCATTTGCGGAGCGCGCGAGGCTGGGCGCTGATGAAGAGATCGAGCGCGGGG 360
Qy 641 GAGGCGCTGCAGAGCTGCACGAGGTGTTTCGCGCGCGCGCTCCCGAGCGCACCGG 700
Db 361 GAGGCGCTGCAGAGCTGCACGAGGTGTTTCGCGCGCGCGCTCCCGAGCGCACCGG 420
Qy 701 CAGAGCTCGGCGAGTGGAGTGCAGCGCACTCGCTGGTCTCGCTTTGTGTGGCG 760
Db 421 CAGAGCTCGGCGAGTGGAGTGCAGCGCACTCGCTGGTCTCGCTTTGTGTGGCG 480
Qy 761 ATCGTGTCCAGCGCCGAGTGGTTCGTGGCGCTGGACAGGCTGGACCGGGGAC 820
Db 481 ATCGTGTCCAGCGCCGAGTGGTTCGTGGCGCTGGACAGGCTGGACCGGGGAC 540
Qy 821 CGTTGGCGGAACAGCGCGCTGGACCTGTACCCCTACGACGCGCGGACGACAGCGCG 880
Db 541 CGTTGGCGGAACAGCGCGCTGGACCTGTACCCCTACGACGCGCGGACGACAGCGCG 600
Qy 881 TTCACCTTCTCTCCCAACTTCGCCACCATTCGCCAGGACACGCGTACCGAGATAAG 940
Db 601 TTCACCTTCTCTCCCAACTTCGCCACCATTCGCCAGGACACGCGTACCGAGATAAG 660
Qy 941 TCCTCTCTCCAGCGCCGCGCACTCTTCTACTACCGCGGCTGAAGGCCCTGCCT 1000
Db 661 TCCTCTCTCCAGCGCCGCGCACTCTTCTACTACCGCGGCTGAAGGCCCTGCCT 720
Qy 1001 CCCATCGCAGGCTGACACTGTGCGGCTGCGACAGAGCCCGAGGCGCTTCATCCCTCC 1060
Db 721 CCCATCGCAGGCTGACACTGTGCGGCTGCGACAGAGCCCGAGGCGCTTCATCCCTCC 780
Qy 1061 GCGCCAGTCTCTCCCGAGGAGCAATGAGATTGTAGACAGCGCTTCAGTTCAGAAAG 1120
Db 781 GCGCCAGTCTCTCCCGAGGAGCAATGAGATTGTAGACAGCGCTTCAGTTCAGAAAG 840
Qy 1121 CGGCTGGAGTGGAGGTCCTCTGCTGCTGCGGACTGTGCGGAGCCACCTGTGG 1180
Db 841 CGGCTGGAGTGGAGGTCCTCTGCTGCTGCGGACTGTGCGGAGCCACCTGTGG 900
Qy 1181 AGGCTCGGACCAAGAGAGGAGTCTGCTACGTCGCGGCTCCAGCCGCCCAACAGCGGAG 1240
Db 901 AGGCTCGGACCAAGAGAGGAGTCTGCTACGTCGCGGCTCCAGCCGCCCAACAGCGGAG 960
Qy 1241 CCCTGCGCGAGCTCGAAGAGAGGCTGAGTGGTCCCTGATACCTGCTCTAAGACCAG 1300
Db 1241 CCCTGCGCGAGCTCGAAGAGAGGCTGAGTGGTCCCTGATACCTGCTCTAAGACCAG 1300
```

[illegible]

Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. .2820

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="primary human renal epithelial cells"

/clone="HRC08032"

/clone_lib="HRC"

/note="cloning vector pME18SFL3"

1. .2820

/note="highly similar to AB027466 Homo sapiens SPON2 mRNA for spordin 2"

535 a 849 c 873 g 563 t

misc_feature

1. .2820

535 a 849 c 873 g 563 t

ORIGIN

Query Match

Best Local Similarity 39.6%; Score 729.2; DB 9; Length 2820;

Matches 731; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1107 CAGTTCAGAAAGCGCGCTGGAGTCCGAGGTCTCCCTGTGGTCTCTGGGACTGTGG 1166
|||||
DB 2079 CAGTTCAGAAAGCGCGCTGGAGTCCGAGGTCTCCCTGTGGTCTCTGGGACTGTGG 2138
|||||
QY 1167 GAGGCGACTGTGGAGGCTCGGACCAAGAGCAGGACTCGTACGTCGGGTCCACCGG 1226
|||||
DB 2139 GAGGCGACTGTGGAGGCTCGGACCAAGAGCAGGACTCGTACGTCGGGTCCACCGG 2198
|||||
QY 1227 CCAACAACGGAGCGCCCTGCCCGAGCTCGAAGAAGAGGCTGAGTCGCTCCCTGATAACT 1286
|||||
DB 2199 CCAACAACGGAGCGCCCTGCCCGAGCTCGAAGAAGAGGCTGAGTCGCTCCCTGATAACT 2258
|||||
QY 1287 GCCTCTAAGACAGAGCGCGGAGCCCTGGGGCCCCCGGAGCCATGGGGTGTCCGGGG 1346
|||||
DB 2259 GCCTCTAAGACAGAGCGCGGAGCCCTGGGGCCCCCGGAGCCATGGGGTGTCCGGGG 2318
|||||
QY 1347 CTCTGTGAGGCTCATGTCTGAGCGCGCGGAGGCGACAGGGGGTTTCGCGCTGCTCCCTG 1406
|||||
DB 2319 CTCTGTGAGGCTCATGTCTGAGCGCGCGGAGGCGACAGGGGGTTTCGCGCTGCTCCCTG 2378
|||||
QY 1407 ACCGCGGTGAGCGCGCGGAGCCATCTGCACTGAAGGGCCCTCTGGTGGCGGCGACGG 1466
|||||
DB 2379 ACCGCGGTGAGCGCGCGGAGCCATCTGCACTGAAGGGCCCTCTGGTGGCGGCGACGG 2438
|||||
QY 1467 GCATTGGGAACAGCCTCTCTCTTCCCAACCTGTCTTAGGGGCCCCCGTGTCCCGT 1526
|||||
DB 2439 GCATTGGGAACAGCCTCTCTCTTCCCAACCTGTCTTAGGGGCCCCCGTGTCCCGT 2498
|||||
QY 1527 CTGCTCTCAGCCTCTCTCTCTCGAGGATAAGTATCTCCCAAGGCTCCAGCTACTCTAA 1586
|||||
DB 2499 CTGCTCTCAGCCTCTCTCTCTCGAGGATAAGTATCTCCCAAGGCTCCAGCTACTCTAA 2558
|||||
QY 1587 ATTATGTCCTCTTATAAGTATTGCTGCTCCAGAGATTGTCTTCATCGTCAGGGGCC 1646
|||||
DB 2559 ATTATGTCCTCTTATAAGTATTGCTGCTCCAGAGATTGTCTTCATCGTCAGGGGCC 2618
|||||
QY 1647 TGCTCCCGAGGTGTCAGATACCTTCAGACCTGGTCTCTAGGCTGTCTGAGGCCACT 1706
|||||
DB 2619 TAGCTCCCGAGGTGTCAGATACCTTCAGACCTGGTCTCTAGGCTGTCTGAGGCCACT 2678
|||||
QY 1707 CTCCCGAGGGCGCATCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTTCGG 1766
|||||
DB 2679 CTCCCGAGGGCGCATCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTTCGG 2738
|||||
QY 1767 AAGCGTCAGTGTTCCTCATGTTATGAGATCTCTCTGCGGTTTGAATAAAGACTATCTCTGTTG 1826
|||||

Db 2739 AAGCGTCAGTGTTCCTCATGTATGGATCTCTCGCTTGAATAAAGACTATCTCTGTTG 2798
|||||
QY 1827 CTCAAAAAATAAAA 1840
|||||
Db 2799 CTCACAAAAAATAAAA 2812
|||||

RESULT 8

AC092535/c

LOCUS

DEFINITION

AC092535

ACCESSION

VERSION

AC092535.4

GI:18640695

KEYWORDS

HTG.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 190462)

Sulston, J.E. and Waterston, R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

99063792

REFERENCE

2 (bases 1 to 190462)

McLellan, M., Cotton, M. and Dixon, R.

The sequence of Homo sapiens BAC clone RP11-20120

Unpublished (2001)

JOURNAL

REFERENCE

3 (bases 1 to 190462)

Waterston, R.H.

Direct Submission

Submitted (16-JUL-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

4 (bases 1 to 190462)

Waterston, R.

Direct Submission

Submitted (08-FEB-2002) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Feb 8, 2002 this sequence version replaced gi:15825624.

Center

Genome Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc

Contact: sapiens@watson.wustl.edu

Summary Statistics

Center project name: H_NH0020120

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.

McPherson, Department of Genetics, Washington University, St. Louis

MO. For additional information about the map position of this

sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male

donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,

Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved

approach for construction of bacterial artificial chromosome

libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-460119; the clone sequenced to the right is RP11-386115. Actual start of this clone is at base position 1 of RP11-20120; actual end is at base position 190462 of RP11-20120.

Data from AC078852 was used to finish this clone, AC092535.

Polymorphisms have been identified between AC078852 and AC092535. Single subclone region exists between 13867 and 13921. There is an unresolved tandem repeat from base 172834 to 175299.

FEATURES

```

source
    Location/Qualifiers
        i. .190462
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosomes="4"
        /map="4"
        /clone="RP11-20120"
        /clone_lib="RPC1-11"
        i. .194
        /rpt_family="L1"
        484. .632
        /note="match to EST AW195334 (NID:g6474404) xn37b05.x1"
        484. .1063
        /note="match to EST AW796230 (NID:g7848100)"
        670. .1086
        /note="similar to Homo sapiens EST BG215528 (NID:g13741549)"
        670. .1086
        /note="similar to Homo sapiens EST BG212839 (NID:g13734526)"
        677. .1159
        /note="similar to Homo sapiens EST BG676306 (NID:g13907702)"
        730. .1142
        /note="similar to Mus musculus EST W57076 (NID:g1358934) mq53d10.rl"
        762. .1159
        /note="similar to Homo sapiens EST BI871289 (NID:g16044964)"
        865. .1142
        /note="similar to Mus musculus EST BF100289 (NID:g10882815)"
        879. .908
        /note="similar to Homo sapiens EST BG260717 (NID:g12770533)"
        930. .1142
        /note="similar to Mus musculus EST AA146274 (NID:g1715663) mq66a10.rl"
        976. .1159
        /note="similar to Homo sapiens EST BG260717 (NID:g12770533)"
        985. .1067
        /note="similar to Homo sapiens EST BG675010 (NID:g13906406)"
        1171. .1197
        /rpt_family="(T)n"
        1195. .1640
        /note="match to EST AA994389 (NID:g3180934) oul0f10.s1"
        1459. .1591
        /note="similar to Mus musculus EST W57076 (NID:g1358934) mq53d10.rl"
        1459. .1804
        /note="similar to Mus musculus EST BF100289 (NID:g10882815)"
        1459. .1871
        /note="similar to Mus musculus EST AA146274 (NID:g1715663) mq66a10.rl"

```

```

misc_feature 1474. .1921
              /note="similar to Homo sapiens EST AW886544 (NID:g8048556)"
misc_feature 1474. .2103
              /note="similar to Homo sapiens EST BG675010 (NID:g13906406)"
misc_feature 1481. .1887
              /note="similar to Homo sapiens EST BG260717 (NID:g12770533)"
misc_feature 1481. .1867
              /note="similar to Homo sapiens EST BI871289 (NID:g16044964)"
misc_feature 1481. .1561
              /note="similar to Homo sapiens EST BG676306 (NID:g13907702)"
misc_feature 1487. .2055
              /note="similar to Mus musculus EST AA049149 (NID:g1528837) mj44cl0.rl"
misc_feature 1500. .2216
              /note="similar to Homo sapiens EST BI838168 (NID:g15949718)"
misc_feature 1575. .1999
              /note="similar to Mus musculus EST AA450532 (NID:g2164202) v954h04.rl"
misc_feature 1577. .1894
              /note="similar to Homo sapiens EST AI934447 (NID:g5673317) wp05h06.x1"
misc_feature 1727. .2209
              /note="similar to Sus scrofa EST BI185352 (NID:g14659761)"
misc_feature 1826. .2338
              /note="similar to Homo sapiens EST AU135255 (NID:g10995794)"
misc_feature 1897. .1987
              /note="similar to Homo sapiens EST AV708740 (NID:g10726005)"
misc_feature 1930. .2380
              /note="match to EST AW296412 (NID:g6703048)"
misc_feature 1939
              /note="similar to Homo sapiens EST BI871289 (NID:g16044964)"
misc_feature 2003. .2338
              /note="similar to Homo sapiens EST AL600859 (NID:g15164365)"
misc_feature 2044. .2330
              /note="similar to Homo sapiens EST AV708740 (NID:g10726005)"
misc_feature 2115. .2338
              /note="similar to Homo sapiens EST BE891093 (NID:g10350077)"
misc_feature 2399. .2480
              /note="similar to Homo sapiens EST AL600859 (NID:g15164365)"
misc_feature 2399. .2598
              /note="similar to Homo sapiens EST BE891093 (NID:g10350077)"
misc_feature 2399. .2482
              /note="similar to Homo sapiens EST AU135255 (NID:g10995794)"
repeat_region 2601. .2687
              /rpt_family="L1"
repeat_region 2710. .3031
              /rpt_family="L1"
repeat_region 3032. .3269
              /rpt_family="MER2_type"

```

Query Match 39.5%; Score 726.4; DB 9; Length 190462;
 Best Local Similarity 99.9%; Pred. No. 9.3e-98;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1107 CAGTTCCAGAAACGCCCTGACTGCGAGGTCTCCCTGTGGTCTCTCTGGGACTGTGCG 1166
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 68431 CAGTTCAGAAACGCCCTGACTGCGAGGTCTCCCTGTGGTCTCTCTGGGACTGTGCG 68372

QY 1167 GAGGCCACTGTGGAGGCTCGGGACCAAGAGCAGGAGTCTGCTACGTCCTCGGTCACGCCG 1226
Db 68371 GAGGCCACTGTGGAGGCTCGGGACCAAGAGCAGGAGTCTGCTACGTCCTCGGTCACGCCG 68312
QY 1227 CCAACAACGGGAGCCCTCGCCCGAGCTCGAAGAAGAGCTGAGTCGCTCCCTGATAACT 1286
Db 68311 CCAACAACGGGAGCCCTCGCCCGAGCTCGAAGAAGAGCTGAGTCGCTCCCTGATAACT 68252
QY 1287 GCCTCTAAGACCAAGAGCCCGAGCCCTCGGGCCCGCCCGAGGCCATGGGTGTCGGGGG 1346
Db 68251 GCCTCTAAGACCAAGAGCCCGAGCCCTCGGGCCCGCCCGAGGCCATGGGTGTCGGGGG 68192
QY 1347 CTCTGTGAGGCTCATGCTGACGGCGCGAGGACAGGGGGTTTCGCGTGCCTCTG 1406
Db 68191 CTCTGTGAGGCTCATGCTGACGGCGCGAGGACAGGGGGTTTCGCGTGCCTCTG 68132
QY 1407 ACCGCGGTGAGGCGCGGCGGACCATCTCTGCACTGAAGGCCCTCTGTTGGCGGCACGG 1466
Db 68131 ACCGCGGTGAGGCGCGGCGGACCATCTCTGCACTGAAGGCCCTCTGTTGGCGGCACGG 68072
QY 1467 GCATTGGGAAACAGCCTCTCTCTTCCCAACCTTGCTTCTAGGGGCCCGCGTGTCCCGT 1526
Db 68071 GCATTGGGAAACAGCCTCTCTCTTCCCAACCTTGCTTCTAGGGGCCCGCGTGTCCCGT 68012
QY 1527 CTGCTCTCAGCCTCCCTCTCTGAGGATTAAGTATCATCCCAAGGCTCAGCTACTCTAA 1586
Db 68011 CTGCTCTCAGCCTCCCTCTCTGAGGATTAAGTATCATCCCAAGGCTCAGCTACTCTAA 67952
QY 1587 ATTATGCTCTTATAGTATTGCTGCTCAGGAGATTTGCTTCATCGTCCAGGGGCC 1646
Db 67951 ATTATGCTCTTATAGTATTGCTGCTCAGGAGATTTGCTTCATCGTCCAGGGGCC 67892
QY 1647 TGCTCTCCAGCTGTGTCAGATACCTCAGACCTGGTGTCTAGGCTGCTGAGGCCACT 1706
Db 67891 TGCTCTCCAGCTGTGTCAGATACCTCAGACCTGGTGTCTAGGCTGCTGAGGCCACT 67832
QY 1707 CTCCGAGGGCGCATCCAGCGGGGCCACTTGAGAAGTGAATAATGGGCGGTTTCGG 1766
Db 67831 CTCCGAGGGCGCATCCAGCGGGGCCACTTGAGAAGTGAATAATGGGCGGTTTCGG 67772
QY 1767 AGCGTCAGTGTTCCTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1826
Db 67771 AGCGTCAGTGTTCCTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 67712
QY 1827 CTCAAAAA 1834
Db 67711 CTCACAAA 67704

RESULT 9
AC022763/c
LOCUS Homo sapiens clone RP11-296G16, WORKING DRAFT SEQUENCE, 24
DEFINITION AC022763
ACCESSION AC022763
VERSION AC022763.2 GI:7652738
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 197176)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-296G16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197176)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Bozulavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced gi:6987625.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5843
Center clone name: 296_G_16
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 185565 bases at least Q40
Consensus quality: 190999 bases at least Q30
Consensus quality: 192979 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 194876; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 438: contig of 438 bp in length
* 439 538: gap of 100 bp
* 539 2494: contig of 1956 bp in length
* 2495 2594: gap of 100 bp
* 2595 3740: contig of 1146 bp in length
* 3741 3840: gap of 100 bp
* 3841 6181: contig of 2341 bp in length
* 6182 6281: gap of 100 bp
* 6282 8921: contig of 2640 bp in length
* 8922 9021: gap of 100 bp
* 9022 12132: contig of 3111 bp in length
* 12133 12232: gap of 100 bp
* 12233 15557: contig of 3325 bp in length
* 15558 15657: gap of 100 bp
* 15658 19986: contig of 4329 bp in length
* 19987 20086: gap of 100 bp
* 20087 21983: contig of 1897 bp in length
* 21984 22083: gap of 100 bp
* 22084 26850: contig of 4767 bp in length
* 26851 26950: gap of 100 bp
* 26951 33137: contig of 4367 bp in length
* 33138 33417: gap of 100 bp
* 33418 36013: contig of 4596 bp in length
* 36014 36113: gap of 100 bp
* 36114 41819: contig of 5706 bp in length
* 41820 41919: gap of 100 bp
* 41920 46044: contig of 4125 bp in length
* 46045 46144: gap of 100 bp

TITLE
JOURNAL
COMMENT

* 46145 52849: contig of 6705 bp in length
* 52850 52949: gap of 100 bp
* 52950 63466: contig of 10517 bp in length
* 63467 63566: gap of 100 bp
* 63567 70466: contig of 6900 bp in length
* 70467 70566: gap of 100 bp
* 70567 80411: contig of 9845 bp in length
* 80412 80511: gap of 100 bp
* 80512 92230: contig of 11719 bp in length
* 92231 92330: gap of 100 bp
* 92331 105063: contig of 12733 bp in length
* 105064 105163: gap of 100 bp
* 105164 115199: contig of 10036 bp in length
* 115200 115299: gap of 100 bp
* 115300 127901: contig of 12602 bp in length
* 127902 128001: gap of 100 bp
* 128002 145911: contig of 17910 bp in length
* 145912 146011: gap of 100 bp
* 146012 197176: contig of 51165 bp in length.

FEATURES

source
1. 197176
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-296G16"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. 438
/note="assembly_fragment"
clone_end:17
vector_side:left
539..2494

misc_feature

/note="assembly_fragment"

misc_feature

2595..3740
/note="assembly_fragment"

misc_feature

3841..6181
/note="assembly_fragment"

misc_feature

6282..8921
/note="assembly_fragment"

misc_feature

9022..12132
/note="assembly_fragment"

misc_feature

12233..15557
/note="assembly_fragment"

misc_feature

15658..19986
/note="assembly_fragment"

misc_feature

20087..21983
/note="assembly_fragment"

misc_feature

clone_end:SP6
vector_side:left

misc_feature

22084..26850
/note="assembly_fragment"

misc_feature

26951..31317
/note="assembly_fragment"

misc_feature

31418..36013
/note="assembly_fragment"

misc_feature

36114..41819
/note="assembly_fragment"

misc_feature

41920..46044
/note="assembly_fragment"

misc_feature

46145..52849
/note="assembly_fragment"

misc_feature

52950..63466
/note="assembly_fragment"

misc_feature

63567..70466
/note="assembly_fragment"

misc_feature

70567..80411
/note="assembly_fragment"

misc_feature

80512..92230
/note="assembly_fragment"

misc_feature

92331..105063
/note="assembly_fragment"

misc_feature

105164..115199
/note="assembly_fragment"

misc_feature

115300..127901
/note="assembly_fragment"

misc_feature 128002..145911
/note="assembly_fragment"
misc_feature 146012..197176
/note="assembly_fragment"
BASE COUNT 46883 a 49476 c 49808 g 48706 t 2303 others
ORIGIN

Query Match 39.5%; Score 726.4; DB 2; Length 197176;
Best Local Similarity 99.9%; Pred. No. 9.3e-98;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1107 CAGTTCAGAAACGCGCTGGACTGCGAGGTCTCCCTGTGGTCTCTGGGACTGTGCG 1166
|||||
Db 129505 CAGTTCAGAAACGCGCTGGACTGCGAGGTCTCCCTGTGGTCTCTGGGACTGTGCG 129446
QY 1167 GAGGCCACTGTGGAGGCTCGGGACCAAGACGAGGACTCGCTACGTCGCGGGTCCAGCCCG 1226
|||||
Db 129445 GAGGCCACTGTGGAGGCTCGGGACCAAGACGAGGACTCGCTACGTCGCGGGTCCAGCCCG 129386
QY 1227 CCAACAACGGAGCCCTGCCCCGAGCTCGAAAGAGAGCTGAGTGCCTCCCTGATAACT 1286
|||||
Db 129385 CCAACAACGGAGCCCTGCCCCGAGCTCGAAAGAGAGCTGAGTGCCTCCCTGATAACT 129326
QY 1287 GCGTCTAAGACCAAGCCCGCAGCCCTTGGGGCCCCCGGAGCCCATGGGGTGTGCGGGG 1346
|||||
Db 129325 GCGTCTAAGACCAAGCCCGCAGCCCTTGGGGCCCCCGGAGCCCATGGGGTGTGCGGGG 129266
QY 1347 CTCCTGTGCAGGCTCATGTCTGAGGGCGCGGAGGGGACAGAGGGGTTTCGGCTGCTCCTG 1406
|||||
Db 129265 CTCCTGTGCAGGCTCATGTCTGAGGGCGCGGAGGGGACAGAGGGGTTTCGGCTGCTCCTG 129206
QY 1407 ACCGGGTGAGCGCGCGGACCATCTCTGCACTGAAGGCCCTCTGTGGCGGCGCACGG 1466
|||||
Db 129205 ACCGGGTGAGCGCGCGGACCATCTCTGCACTGAAGGCCCTCTGTGGCGGCGCACGG 129146
QY 1467 GCATTGGGAAACAGCCCTCTCTTCCCAACCTTGCTTTAGGGGCCCCCCCGTGTCCCGT 1526
|||||
Db 129145 GCATTGGGAAACAGCCCTCTCTTCCCAACCTTGCTTTAGGGGCCCCCCCGTGTCCCGT 129086
QY 1527 CTGCTCTCAGCCTCTCTCTGCGAGGATAAGTCAATCCCAAGGCTCCAGCTACTCTAA 1586
Db 129085 CTGCTCTCAGCCTCTCTCTGCGAGGATAAGTCAATCCCAAGGCTCCAGCTACTCTAA 129026
QY 1587 ATTATGCTCTCTATTAAGTTATTGTCTCCAGGAGATTGCTTCTCATCGTCCAGGGGCC 1646
|||||
Db 129025 ATTATGCTCTCTATTAAGTTATTGTCTCCAGGAGATTGCTTCTCATCGTCCAGGGGCC 128966
QY 1647 TGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGTCTTAGGCTGTGCTGAGCCCACT 1706
|||||
Db 128965 TGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGTCTTAGGCTGTGCTGAGCCCACT 128905
QY 1707 CTCCTCCAGGGCGCATCCAAAGCGGGCGGCGCTTGAGAGTGAATAAATGGGCGGTTTCGG 1766
|||||
Db 128905 CTCCTCCAGGGCGCATCCAAAGCGGGCGGCGCTTGAGAGTGAATAAATGGGCGGTTTCGG 128846
QY 1767 AAGCGTCAGTGTTCATGATCTCTCTGCGCTTTGAATAAAGACTATCTCTGTTG 1826
|||||
Db 128845 AAGCGTCAGTGTTCATGATCTCTCTGCGCTTTGAATAAAGACTATCTCTGTTG 128786
QY 1827 CTCAAAAA 1834
|||||
Db 128785 CTCACAAA 128778

RESULT 10
AK024499
LOCUS Homo sapiens mRNA for FLJ00108 protein, partial cds.
DEFINITION AK024499
ACCESSION AK024499
VERSION AK024499.1 GI:10440517
KEYWORDS fis (full insert sequence).
SOURCE Homo sapiens adult spleen cdna to mRNA, clone:as00108.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4292)
Obara, O., Nagase, T., Kikuno, R. and Okumura, K.
The nucleotide sequence of a long cDNA clone isolated from human
spleen
Published Only in Database (2000) In press
2 (bases 1 to 4292)
Obara, O., Nagase, T., Kikuno, R. and Okumura, K.
Direct Submission
Submitted (24-AUG-2000) Osamu Obara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaifo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/NEDO, Tel: 81-438-52-3913,
Fax: 81-438-52-3914)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Kazusa DNA Research Institute.
FEATURES
Location/Qualifiers
1..4292
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="as00108"
/tissue_type="spleen"
/dev_stage="adult"
/note="vector:pBluescriptII SK plus"
1..653
/gene="FLJ00108"
1..653
/gene="FLJ00108"
/note="For this clone, Genemark analysis triggered an
alert for spurious CDS split. The result of Genemark
analysis of the nucleotide sequence of this clone is
accessible through http://www.kazusa.or.jp/NEDO.
Start codon is not identified."
/codon_start=3
/product="FLJ00108 protein"
/protein_id="BAB15789.1"
/db_xref="GI:10440518"
/translation="SAELVQRHRLSVFVVRVPSDFVGVDSLDLGDGRWREQA
ALDLYPDAGTDSGFSPNPFATIPDVTVEVSHGACAAAPAVGTGLGPGCTL
ASQPTAACAGFVEALLPAVSKGVQVSTLTPRHSTAGTAGRASRGRLPLAS
VSSPCSVGLRWGHCTLSPTVVVRVWTLGSGPSAIRIPEGLLGPQANLFLPR"
BASE COUNT 788 a 1323 c 1349 g 832 t
ORIGIN
Query Match 39.3%; Score 724; DB 9; Length 4292;
Best Local Similarity 100.0%; Pred. No. 4.7e-97;
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1107 CAGTTCAGAAAGCGCTGGAGTGGAGGTCTCCCTGTGGTCTCTGGGACTGTGGG 1166
|||||
DB 3568 CAGTTCAGAAAGCGCTGGAGTGGAGGTCTCCCTGTGGTCTCTGGGACTGTGGG 3627
|||||
QY 1167 GAGGCACATCTGGAGGCTCGGACCAAGCAGGAGTCTGCTACGTCGCGGTCCAGCCCG 1226
|||||
DB 3628 GAGGCACATCTGGAGGCTCGGACCAAGCAGGAGTCTGCTACGTCGCGGTCCAGCCCG 3687
|||||
QY 1227 CCAACAACGGAGCCCTGCCCGAGCTCGAAGAAGAGGCTGAGTGCCTGTGATAACT 1286
|||||
DB 3688 CCAACAACGGAGCCCTGCCCGAGCTCGAAGAAGAGGCTGAGTGCCTGTGATAACT 3747
|||||
QY 1287 GCGTCTAAGACAGAGCCCGAGCCCTGGGGCCCCCGGAGCCATGGGTGTTCGGGGG 1346
|||||
DB 3748 GCGTCTAAGACAGAGCCCGAGCCCTGGGGCCCCCGGAGCCATGGGTGTTCGGGGG 3807
|||||
QY 1347 CTCCTGTGAGGCTCATGTGTCAGGCGCGGAGGGCACAGGGGGTTTCGCGTGTCTCTG 1406
|||||
DB 3808 CTCCTGTGAGGCTCATGTGTCAGGCGCGGAGGGCACAGGGGGTTTCGCGTGTCTCTG 3867
|||||
QY 1407 ACCGCGGTGAGCGCGCGCCGACCATCTCTGCACTGAAGGGCCCTCTGTGTGGCGGCACGCG 1466
|||||

Db 3868 ACCGCGGTGAGCGCGCGGACCATCTCTGCACTGAAGGGCCCTCTGTGTGGCGGCACGCG 3927
|||||
QY 1467 GCATTGGGAACAGAGCTCTCTCTTTCCCAACCTTCTCTTTAGGGGCGCCCGTGTCCCGT 1526
|||||
Db 3928 GCATTGGGAACAGAGCTCTCTCTTTCCCAACCTTCTCTTTAGGGGCGCCCGTGTCCCGT 3987
|||||
QY 1527 CTGCTCTCAGCCTCTCTCTCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCTAA 1586
|||||
Db 3988 CTGCTCTCAGCCTCTCTCTCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCTAA 4047
|||||
QY 1587 ATTATGCTCTCTCTCTCTCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCTCTAA 1646
|||||
Db 4048 ATTATGCTCTCTCTCTCTCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCTCTCTCT 4107
|||||
QY 1647 TGGCTCTCCACGCTGTTGTCAGATACCTCAGACCTGCTAGGCTGTCTGAGCCCTACT 1706
|||||
Db 4108 TGGCTCTCCACGCTGTTGTCAGATACCTCAGACCTGCTAGGCTGTCTGAGCCCTACT 4167
|||||
QY 1707 CTCGCGAGGGCGCATCCCAAGCGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTTCGG 1766
|||||
Db 4168 CTCGCGAGGGCGCATCCCAAGCGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTTCGG 4227
|||||
QY 1767 AAGCGTCAAGTGTTCATGTTATGATGATCTCTCTGTTGTAATAAGACTATCTCTGTGG 1826
|||||
Db 4228 AAGCGTCAAGTGTTCATGTTATGATGATCTCTCTGTTGTAATAAGACTATCTCTGTGG 4287
|||||
QY 1827 CTCGA 1830
|||||
Db 4288 CTCGA 4291
|||||
RESULT 11
AF155196 AF155196 2325 bp mRNA linear ROD 23-DEC-1999
LOCUS Rattus norvegicus mindin precursor, mRNA, complete cds.
DEFINITION AF155196
ACCESSION AF155196
VERSION AF155196.1 GI:5031505
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2325)
AUTHORS Feinstein, Y., Borrell, V., Garcia, C., Burstin-Cohen, T., Tzarfaty, V.,
Frumkin, A., Nose, A., Okamoto, H., Higashijima, S., Soriano, A. and
Klar, A.
TITLE F-spondin and mindin: two structurally and functionally related
genes expressed in the hippocampus that promote outgrowth of
embryonic hippocampal neurons
JOURNAL Development 126 (16), 3637-3648 (1999)
MEDLINE 99339921
PUBMED 10409509
REFERENCE 2 (bases 1 to 2325)
AUTHORS Feinstein, Y., Borrell, V., Garcia, C., Burstin-Cohen, T., Tzarfaty, V.,
Frumkin, A., Nose, A., Okamoto, H., Higashijima, S., Soriano, A. and
Klar, A.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1999) Anatomy & Cell Biology, Hebrew University,
PO Box 12272, Jerusalem 91120, Israel
FEATURES
Location/Qualifiers
1..2325
/organism="Rattus norvegicus"
/strain="Sprague Dawley"
/db_xref="taxon:10116"
512..577
512..1504
/note="adhesion molecule"
/codon_start=1
/product="mindin precursor"
/protein_id="AAD38195.1"
/db_xref="GI:5031506"
sig_peptide
CDS


```
/db_xref="taxon:10090"
/clone="MGC:27697 IMAGE:4923292"
/tissue.type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_S62"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
228..1220
CDS
/codon_start=1
/product="Similar to spondin 2, extracellular matrix protein"
/protein_id="AAH17616.1"
/db_xref="GI:17160895"
/translation="MENISLALRALWFLVLLAMIGSTTSQPLGSESVCTARPLARYSI
TGIGKWSQTFAPKQYPLFRPPAQWSSLLGAHSSDYSMWRKNHYSVNGLRDPAERGEA
WLMKEIEAGKQSLQVAVFSAPALPSTGOTSTLELVHPRHSLVSVFVRIVPSPDW
FVGLDLCEGGRWKEQVLDLYPHDAGTDSGTFPSPNFATIFQDVTETASPS
HPANFYPYPLSLPIAKTVFRIHQSPRAFPPSLDLASRNEIIVSLSPVETPLD
CEVLSWSSMLGCGPCGKLGAKSRTRYVRVPANNGTFCPELEEEAEAPDNCV"
BASE COUNT 410 a 503 c 486 g 379 t
ORIGIN
Query Match 36.9%; Score 679; DB 10; Length 1778;
Best Local Similarity 75.5%; Pred. No. 2.2e-90;
Matches 928; Conservative 0; Mismatches 280; Indels 21; Gaps 6;
QY 94 CGCCGCGCCCGCAGCCCTTCTCCCTCTTCTCCACGCTCTCTATCTGCTCTCTCGCTGGA 153
DB 19 CCCTCGCAGCAGGAACCTCTTGCTCTCTCTCAAGCCCTCTCGCAGCTTCTCGG- 77
QY 154 GGCAGGCGCTGCAGCATCGAAGACAGGAGAACTGGAGCTCATTTGGCGCGCCGGGGC 213
DB 78 GACCAGGCACTGAGGCA--GAAGACAGGAGAGCTGGGGCT-ATCTGCTGGCCAGAGGC 134
QY 214 GCCGCGCTCGGCTTAAATAGGAGCTCCGGGCTCTGCTGGGACCGACGCTGCCGGCC 273
DB 135 ACTAACCTTGGGGCACATAGGATCTGGGAGCTCTGGGACCAATCTCCAGCTGGAACG 194
QY 274 GC-----GCTCCGCGCTCTCTCGCGGGTGATGAAACCCAGCCGCGCGCGCC 325
DB 195 CCAGCAGCCTGCCTCTGACGCTTTTGCCAGGTGATGAAACCTTGAGTCTT-----GCC 248
QY 326 CTGGGAAGCCCTCTGCGCTCTCTCTCTGCGCCACTCTGCGCGCCGCC---GGCCAGCCT 382
DB 249 CTGGGAGAGCTCTTTGGGTCTTCTCTCTCTGCGCATGATAGCTCCACACGAGCCAGCG 308
QY 383 CTTTGGGGAGAGTCCATCTGTTCGCCAGAGCCCGGCCCAATACAGCATCACCTTCACG 442
DB 309 CTGGGGGAGAAATCGGTTGTACAGCCCGGCGCTCTGGCTAGATACAGCATCACTTTCATT 368
QY 443 GGCAAGTGGAGCCAGCGGCTTCCCAAGCAGTACCCCTGTTCGCGCCCGCCCTGCGCAG 502
DB 369 GGCAAGTGGAGCCAGCAGCATTTCCCAAGCAGTACCCCTGTTCGCGCCCGCTGCACAG 428
QY 503 TGTCTTTCGCTGTGGGGCGCGGATAGTCTCGACTACAGCATGTGGAGGAAGAACCCAG 562
DB 429 TGTCTCTCTCTGCTGGGGGAGCTCACAGCTCTGACTACAGCATGTGGGGGAAGATGAG 488
QY 563 TAGCTCAGTAACGGCTGCGCGACTTTGGGAGCGCGCGGAGCCCTGGGCGCTGATGAG 622
DB 489 TATGTACAGCAATGGCTGAGGGACTTTGCTAGCGGTGGTGGAGCCCTGGGCACTGATGAG 548
QY 623 GAGATCAGGCGCGGGGAGGCGCTGCAGAGGCTGCAGAGGCTGCAGAGGTGTTTCGGCGCCCGC 682
DB 549 GAGATCGAAGCTCAGSAGAGAACTTCAGAGTGTGCAGCAGTGTCTCAGCTCTCGCC 608
QY 683 GTCCACAGGCGACCGGGCAGAGCTCGGGAGAGCTGGAGTGCAGCGCAGGCACTCGCTG 742
DB 609 ATCCCTAGCGGTACCGGGCAGACTTCCACAGAGTTAGAGTGCACCTAGGCACTCGCTG 668
QY 743 GTCTCTCTTGTGTCGCTCGTGGCCAGCCCGACTGCTTCTGCGGCTGGACACCTCG 802
DB 669 GTGCTCTCTGTTAGCATTTGCCCGACCCCTGACTGTTGTGGGCATCGACAGCTCTG 728
```

```
QY 803 GACCTGTGCGACGGGACCGCTTTGGCGGGAACAGAGCGCGCTGAGACCTGTACCCCTACGAC 862
DB 729 GACCTGTGTGAGGAGCGCGCTTGAAGAGCAGAGTGGTCTGGATCTTTATCCACAGAT 788
QY 863 GCGGGAGCGACAGCGGCTTTCACCTTCTCTCCCAACTTCCGCCACCATCGCGGAGGAC 922
DB 789 GCAGGGAGCGACAGTGGTTCACCTTCTCTCCCTTAACCTTTGCAACTATCCCAAGAC 848
QY 923 AGGTCACCGAGATACAGTCTCTCTCCAGCCACCGCCGCAACTCTTCTACTACCG 982
DB 849 ACAGTACTGAGATAACAGCTCATCGCCAGCCACCCAGCAATTCATCTACTACCA 908
QY 983 CGGCTGAAGGCGCTGCTCCCTCCGACGAGGTGACACTGCTGGCTGGACAGAGCCCC 1042
DB 909 CGCCTCAAGTCTCTTACCTCCCAAGTACACCTTCTGCGGGCTACAGCATAGTCCC 968
QY 1043 AGGCGCTTCACTCCCTCCCGCCAGTCTCTGCGCAGCAGGACAATGAGATTGTAGACAG 1102
DB 969 AGGGCGCTTTCGCCACCTTTCCTGGACCTGGCCAGCGGCAACCAATCGTTGACAG 1028
QY 1103 GCCTCAGTTCACAGAACGCGCTGGACTGGAGTCTCCCTGCTGGTCTGCTGGGACTG 1162
DB 1029 CTTTCAAGTTCACAGAGACACGCTGGACTGTGAGTGTTCCTGTGGTCACTCTGGGACTG 1088
QY 1163 TCGGAGGCGACTGTGGGAGGCTCGGGACCAAGAGCAGGACTCGCTACGTCGGGTCAG 1222
DB 1089 TGTGAGGAGCATGTGAAAGTTGGGAGCAAGAGCAGAACTCGCTATGTCGCTGTTCA 1148
QY 1223 CCGGCCAACAAGGGAGCCCTGCGCCGAGCTCGAAGAGAGGCTGAGTGCCTGCTGAT 1282
DB 1149 CTTGCTTAACAATGGGACTCTCTGCGGAGCTTGAAGAAGAGCGCGAGTGTGCCAGAT 1208
QY 1283 AACTGCTCTAAGACAGAGCCCGCCGAC 1311
DB 1209 AACTGCTCTAACCAGACTCCAGTAGC 1237
RESULT 13
BD006593
LOCUS Integrin ligand, human Mindin.
DEFINITION BD006593
ACCESSION BD006593.1 GI:18634964
VERSION JP 2001500386-A/2.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 1020)
AUTHORS Jonak,Z.L., Trullis,S.H., Tsui,P. and Lane,P.A.
TITLE Integrin ligand, human Mindin
JOURNAL Patent: JP 2001500386-A 2 16-JAN-2001;
SMITHKLINE BEECHAM CORP
COMMENT OS Homo sapiens (human)
PN JP 2001500386-A/2
PD 16-JAN-2001
PF 07-MAY-1998 JP 1998548555
PR 09-MAY-1997 US 60/046106
PI ZENKA L JONAK STEPHEN H TRULLI, PING TSUI, PAMELA A LANE PC
A61K39/395,A61K48/00,C07H21/04,C07K14/435,C07K16/00, PC
C07K16/28,
PC C12N15/11,C12N15/63,G01N33/53
CC
FH Key Location/Qualifiers
FT source 1..1020
/organism='Homo sapiens (human)'.
FEATURES
Location/Qualifiers
1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 183 a 333 c 306 g 198 t
ORIGIN
```

Query Match		29.9%; Score 549.8; DB 6; Length 1020;
Best Local Similarity		96.2%; Pred. No. 2.1e-71;
Matches		563; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY	1256	GAAGAAGAGCTGAGTGGCTCCCTGATTAAGTGGCTTAAGACACAGAGCCCGCAGCCCTT 1315
Db	430	GAACGGCGGAGCCCTGGCGCTGATTAAGAGAGATCTAAGACACAGAGCCCGCAGCCCT 489
QY	1316	GGGCCCCCGGAGCATGGGGTGTGGGGGCTCCCTGTGAGGCTCATGCTGCAGCGGC 1375
Db	490	GGGCCCCCGGAGCATGGGGTGTGGGGGCTCCCTGTGAGGCTCATGCTGCAGCGGC 549
QY	1376	CGAGGCACAGGGGTTTCGCGTGTCTCTGACCGGGGTGAGCGCGCGACCATCTCT 1435
Db	550	CGAGGCACAGGGGTTTCGCGTGTCTCTGACCGGGGTGAGCGCGCGACCATCTCT 609
QY	1436	GCATGAAGGGCCCTGTGGTGGCGGACAGGGGATTTGGGAAACAGGCTCTCTTTCCCA 1495
Db	610	GCAATGAAGGGCCCTGTGGTGGCGGACAGGGGATTTGGGAAACAGGCTCTCTTTCCCA 669
QY	1496	ACCTTGCTCTTAGGGGCCCGCGTCCGCTCTGCTCTGACGCTCTCTCTCTCTCTCTCT 1555
Db	670	ACCTTGCTCTTAGGGGCCCGCGTCCGCTCTGCTCTGACGCTCTCTCTCTCTCTCTCT 729
QY	1556	AAAGTCATCCCCAAGGCTCCAGCTACTCTAAATTAATGTCTCTCTCTCTCTCTCTCT 1615
Db	730	AAAGTCATCCCCAAGGCTCCAGCTACTCTAAATTAATGTCTCTCTCTCTCTCTCTCT 789
QY	1616	CAAGGAGATTGCTTTCATCGTCCAGGGGCTGGCTCCCAAGCTGTTGAGATACCTCAG 1675
Db	790	CAAGGAGATTGCTTTCATCGTCCAGGGGCTGGCTCCCAAGCTGTTGAGATACCTCAG 849
QY	1676	ACCTGTGCTCTAGGCTGTGCTGAGCCCACTCTCCGAGGGCGCATCCAGCGGGGCCA 1735
Db	850	ACCTGTGCTCTAGGCTGTGCTGAGCCCACTCTCCGAGGGCGCATCCAGCGGGGCCA 909
QY	1736	CTTGAGAAGTGAATAATGGGGCGGTTTCGGAAGCGTCACTGTTTCCATGTTATGGATCT 1795
Db	910	CTTGAGAAGTGAATAATGGGGCGGTTTCGGAAGCGTCACTGTTTCCATGTTATGGATCT 969
QY	1796	CTCTGGGTTTGAATAAGACTATCTCTGTGTGCTCAAAAAA 1840
Db	970	CTCTGGGTTTGAATAAGACTATCTCTGTGTGCTCAAAAAA 1014
RESULT 14		
AB006085		
LOCUS		2116 bp mRNA linear VFT 10-MAY-2001
DEFINITION		Danio rerio mRNA for MINDIN2, complete cds.
ACCESSION		AB006085
VERSION		AB006085.1 GI:2529222
KEYWORDS		MINDIN2.
SOURCE		Danio rerio embryo cDNA to mRNA.
ORGANISM		Danio rerio
REFERENCE		1 (sites)
AUTHORS		Higashijima, S., Nose, A., Eguchi, G., Hotta, Y. and Okamoto, H.
TITLE		Mindin/F-spondin family: novel ECM proteins expressed in the zebrafish embryonic axis
JOURNAL		Dev. Biol. 192 (2), 211-227 (1997)
MEDLINE		98104230
REFERENCE		2 (bases 1 to 2116)
AUTHORS		Higashijima, S.
TITLE		Direct Submission
JOURNAL		Submitted (28-JUL-1997) Shin-ichi Higashijima, National Institute for Basic Biology, Division of Morphogenesis, Nishigonaka 38, Myodaiji-cho, Okazaki, Aichi 444, Japan (E-mail: shinichi@nibb.ac.jp, Tel: 81-564-55-7572, Fax: 81-564-55-7571)
FEATURES		Location/Qualifiers
source		1..2116
		/organism="Danio rerio"

CDS		/db_xref="taxon:7955"
		/dev_stage="embryo"
		659..1654
		/codon_start=1
		/product="MINDIN2"
		/protein_id="BAA22809.1"
		/db_xref="GI:2529223"
		/translation="METMTSLRVNCWLTMTALLSGVPAMPVDVDMCTAPSTAKYRL
		TTTGWOTAFPKHYPLYPAPQWSPGLVTHSSDYHLWQNEYASNGVRESEAREA
		WTLIKEVAAGERIQSYGLFAPAVVAGTGHATTEFEVFAHSLLSFVLRVPSDW
		FIGVDSLNCEGDHGENLSLELYPYDAGTDSGTFSSPNFETIPDQKVTQITSFSP
		HPANFSFYPRLKLPIAKVSLTKIKNNQIFSLPIQTSNQIPSGNEIDGLINTPL
		DEEVSWSPWGLCKGCGEKGVKHTRYTHMHPANNAGPCPSLEEKRLCIPNCV"
BASE COUNT		633 a 478 c 413 g 592 t
ORIGIN		
Query Match		19.8%; Score 364.2; DB 5; Length 2116;
Best Local Similarity		62.1%; Pred. No. 2.7e-44;
Matches		592; Conservative 0; Mismatches 358; Indels 3; Gaps 1;
QY	365	GGCGCGCGCGCCAGCTCTTTGGGGGAGAGTCCATCTGTTCCGCCAGAGCCCGGCCAAA 424
Db	722	GGAGTCCCTGCAATGCGGTGGATGTGACCGCCCTGTGCACGGCCCTTCCACGGCTAAG 781
QY	425	TACAGCATCACCTTTCACGGGCAAGTGAGCCAGACAGCGGCTTCCCAAGAGTACCCCTG 484
Db	782	TACCGACTGACCTTACTGCGCAGTGGACACAGACTGCTTCCCAAGCAGTACCCCTG 841
QY	485	TTCCGCGCCCTCGCGAGTGGTCTTTCGCTGCTGGGGCGCGGATAGCTCCGACTACAGC 544
Db	842	TACAGGCCACTGCTCAGTGGTCTCTCTCTTATTGGAGTAACCTCACAGTTTCAGACTATCAC 901
QY	545	ATGTGGAGGAGAACAGTACGTACGTAAACGGGCTGCGGACCTTTCGGAGCGCGGCGAG 604
Db	902	CTCTGCAAGGAATGAGTACGCCCAATATGGGGTGAGAGAGTCTCTGAGAGAGACAG 961
QY	605	GCCTGGCGCTGATGAAGGAGATCGAGCGCGGGGAGGCGCTGCGAGAGCGTGCACGAG 664
Db	962	GCYTTGGACCTGATAAAGGAAGTGGAGCGAGCTGGGGAACGGATCCAGAGCGTCTATGGC 1021
QY	665	GTGTTTTGGCGCGCGCTCCCGAGCGGACCGGGGAGAGCGTGGCGGGAGCTGGAGGTG 724
Db	1022	CTCTTCTCAGCGCGAGCTGTGGTTGACGAACTGGCCACGCAACACAGAAATTTAGGTC 1081
QY	725	CAGCGAGGCACTCGCTGGTCTGCTTTGTGCTGGCATCTGTCACCCCGGCTGCTGCTTC 784
Db	1082	TTTGGGAGACACTCTTTGCTGCTCTCTTTCATTTGTCGATTTGCGCAAGTCTCTGCTG 1141
QY	785	GTGGGGTGGACAGCTTGGACCTGTGCGACGGGAGCGGTTGGGGGAACAGCGCGCTG 844
Db	1142	ATTGGAGTAGACAGTCTGAACCTTGTGCGAGGCGCACCCAGGTAAAGAAACATAAAGCTA 1201
QY	845	GACCTGTACCCCTACGACCGCGGAGCGAGCGGCTTTCACCTTCTCTCTCCCAACTTC 904
Db	1202	GAGCTCTACCCCTATGATGCGCGCACAGACAGTGGTTTACCTTCTCTCTCGCCAACTTT 1261
QY	905	GCACCATCCCGCAGGACAGGTGACCGAGATACGCTCTCTCTCCAGCCACCCCGGCC 964
Db	1262	GAACCACTTCTCAAGACAAGGCTCACTGACATACCTCATCATTCCTCCCAAGCCACCA 1321
QY	965	AACTCTTCTACTACCGCGGCTGAAGCGCTGCTCCCATCGCGAGGTTGACACTGGTG 1024
Db	1322	AACTCTTCTACTATCCAGATTTGAAGCACTCTCCCGCAATTTGCCAAGTCTCTCTAAC 1381
QY	1025	CGGCTCGCAGAGCCCCAGGGCTTTTCATCCCTCCCGCCCCAGTCTCTGCCAGAGGAC 1084
Db	1382	AAGATCAAGAAACAACAGATCTTCAGCTTACCCTATCCCAATCCAGCAACACAGTCCAATCAGATC 1441
QY	1085	---AATGAGATTGTAGACACGGCTCAGTTCCAGAAACGGCGCTGAGCTCGGAGGTCTCC 1141
Db	1442	CCAAGCGGGAATGAAATAGACGGGCTCTTATAAATACCCCTCTCTGAGCTCGGAGGTCTCT 1501
QY	1142	CTGTGCTGCTCTGGGAGCTGTGGGAGGCACTGTGGAGGCTCTGGGACCAAGAGCAGG 1201

```

||||| | |||| | || | || | |||| | || | || |
Db 1502 GTGTGTCACCTTGGGGTCTTTGTAAAGGTCAGTGTGTGAGAAAGGGTAAACACAGA 1561

||||| | |||| | || | |||| | || | |||| | || | |||| | || |
QY 1202 ACTCGCTACGTCGGGTCACCGCCGCAACACGGGAGCCCTGCCCGAGCTCGAAGAA 1261

||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 1562 ACCGGTATATTACATGCATCTCGAACAACGGTGCTCCCTGTCTCCCTAGAGAA 1621

||||| | |||| | || | |||| | || | |||| | || | |||| | || |
QY 1262 GAGGCTGAGTCGCTGCCGTGATACTAAAGACAGAGCCCGCAGCC 1314

||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 1622 AAGAGACTCTGTATCCCTGACAACTGTGTGAATGTAAATCAGATGAGCCTC 1674

RESULT 15
AR037874
LOCUS
DEFINITION Sequence 1 from patent US 5804382.
ACCESSION AR037874
VERSION AR037874.1 GI:5956591
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 400)
AUTHORS Sytkowski,A.J. and Yang,M.
TITLE Methods for identifying differentially expressed genes and
JOURNAL differences between genomic nucleic acid sequences
FEATURES
Location/Qualifiers
source
1..400
/organism="unknown"
BASE COUNT 70 a 157 c 110 g 63 t
ORIGIN
Query Match 18.7%; Score 343.8; DB 6; Length 400;
Best Local Similarity 97.0%; Pred. No. 3.7e-41;
Matches 393; Conservative 0; Mismatches 7; Indels 5; Gaps 4;

QY 850 GTACCCCTACGACGCGGGACGGACAGCGGCTTACACTTCTCCCTCCCAACTTCGCCAC 909
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 1 GTACCCCTACGACGCGGGAC -GACAGCGGCTTACACTTCTCCCTCCCAACTTCGCCAC 59

QY 910 CATCCCGCAGACACGCTGACCGAGATAGCTCTCTCCAGCCACCCGCCCACTC 969
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 60 CATCCCGCA -GACACGGTGACCGAGATAGCTCTCTCCAGCCACCCGCCCACTC 118

QY 970 CTCTACTACCGCGCTGAAGCCCTGCCCTCCCATCGCAGGCTGACACTGTGGGCT 1029
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 119 CTCTACTACCGCGCTGAAGCC -TGCTCCCATCGCAGGCTGACACTGTGGGCT 176

QY 1030 GCGACAGAGCCCGAGGCGCTTCATCCCTCCCGCCAGTCCTGCCAGCAGGCAATGA 1089
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 177 GCGACAGAGCCCGAGGCGCTTCATCCCTCCCGCCAGTCCTGCCAGCA -GGACAATGA 235

QY 1090 GATTGTAGACAGCGCTCAGTCCAGAAACCGCGCTGGAGTCCCGTGTGGTC 1149
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 236 GATTGTAGACAGCGCTCAGTCCAGAAACCGCGCTGGAGTCCCGTGTGGTC 295

QY 1150 GTCTGGGGACTGTGGGAGGCGACCTGTGGAGGCTCGGACCAAGAGAGGACTCGCTA 1209
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 296 GTCTGGGGACTGTGGGAGGCGACCTGTGGAGGCTCGGAGTCCAGAGAGGACTCCGTA 355

QY 1210 CGTCCGGGTCACCGCGCCCAACACGGGAGCCCTGCCCGGAGCT 1254
||||| | |||| | || | |||| | || | |||| | || | |||| | || |
Db 356 CGCCCGGGTCCAGCCCGCCCAACACGGGAGCCCTGCCCGGAGCT 400
```

Search completed: October 27, 2002, 10:47:17
Job time : 3787 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2002, 03:39:47 ; Search time 298 Seconds
(without alignments)
10601.082 Million cell updates/sec

Title: US-09-807-200-1
Perfect score: 1840
Sequence: 1 ctgcagcggggtgcggcag.....ctgtgtctcaaaaaaaaaa 1840

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

1:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1840	100.0	1840	21	AAZ95034
2	1827.6	99.3	1991	22	AAI58028
3	1825.6	99.2	1926	22	AAI59814
4	1812.8	98.5	1846	20	AAZ34089
5	1812.8	98.5	1847	21	AAC78507
6	1812.8	98.5	1847	21	AAC58630
7	1812.8	98.5	1847	21	AAA49728
8	1793.2	97.5	1798	22	AAK94244
9	1761.2	95.7	1785	22	AAF90566

10	1760.2	95.7	1770	22	AAF90567	Human extracellular
11	1692.4	92.0	1779	22	RAAD20094	Human NPG-1 cDNA.
12	1686.2	91.6	1718	21	AAC59794	Human secreted pro
13	1680.8	91.3	1848	22	RAH34981	Human colon cancer
14	1666.4	90.6	1669	22	AAK94182	Human full-length
15	1563.4	85.0	1607	19	AAV63241	cDNA encoding an a
16	1297.2	70.5	1458	21	AAZ52575	Human secreted pro
17	1100.2	59.8	1105	20	AAZ02019	Human NAF-1 DNA.
18	1098.6	59.7	1105	18	AAZ78360	Human neuronal att
19	988	53.7	996	20	AAV72537	Human mindin cDNA.
20	734	39.9	2261	21	AAC59258	Human secreted pro
21	700.8	38.1	763	22	AAK91675	Human cDNA 5'-end
22	700.8	38.1	763	22	AAK93839	Human cDNA clone r
23	684.2	37.2	993	19	AAV63258	Degenerate sequenc
24	644.6	35.0	825	22	AAK91827	Human cDNA 5'-end
25	644.6	35.0	825	22	AAK93227	Human cDNA clone r
26	547.4	29.8	1021	20	AAV72538	Human mindin-relat
27	483	26.2	534	22	AAK92493	Human cDNA 3'-end
28	343.8	18.7	400	19	AAV53726	Nucleotide sequenc
29	339	18.4	400	20	AAZ06948	Human NPG-1 partia
30	304.8	16.6	506	20	AAZ02024	Human NAF-1 DNA ho
31	301.4	16.4	539	22	AAK87746	Human digestive sy
32	301.4	16.4	539	22	AAI57550	Human colorectal c
33	233.4	12.7	432	19	AAV63270	EST sequence encod
34	217.8	11.8	553	19	AAV63268	EST sequence encod
35	204.2	11.1	316	20	AAZ02026	Human NAF-1 DNA ho
36	204.2	11.1	316	20	AAZ02025	Human NAF-1 DNA ho
37	204	11.1	204	21	AAZ95016	Prostate cancer sp
38	191	10.4	541	19	AAV63266	EST sequence encod
39	163.6	8.9	470	19	AAV63267	EST sequence encod
40	121.4	6.6	2574	23	ABL11611	Drosophila melanog
41	106.6	5.8	2450	23	ABL11997	Drosophila melanog
42	81.6	4.4	2646	23	ABL19249	Drosophila melanog
43	79.6	4.3	4061	20	AAZ30094	Bovine vascular sm
44	77.8	4.2	7273	23	ABL11996	Drosophila melanog
45	75.4	4.1	114955	20	AAZ53491	Human adenosine A1

ALIGNMENTS

RESULT 1
AAZ95034
ID AAZ95034 standard; cDNA; 1840 BP.
AC AAZ95034;
XX
XX
DT 15-AUG-2000 (first entry)
XX
DE Cancer specific gene Prol08 cDNA useful as prostate cancer marker.
XX
KW Cancer specific gene; CSG; prostate cancer; diagnosis; monitoring;
KW staging; imaging; metastasis; therapy; human; Prol08; ss.
OS Homo sapiens.
FH Key
FT CDS
FT 299..1294
/*tag= a
XX
XX
PN WO200023108-A1.
XX
PD 27-APR-2000.
XX
PF 18-OCT-1999; 99WO-US23764.
XX
PR 19-OCT-1998; 98US-0104741.
XX
PA (DIAD-) DIADEXUS LLC.
XX
PI Ali SM, Sun Y, Salceda S, Recipon H, Caferkey R;
XX WPI; 2000-339528/29.
DR

DR P-PSDB: AAY79561.

xx Diagnosing, detecting, staging, monitoring, imaging and treating
 PT cancers, especially useful for detecting prostate cancer comprises
 PT measuring changes in levels of cancer specific genes in cells, tissues
 PT and body fluids
 xx
 xx Claim 6: Page 28-29: 35pp: English.
 PS

The present sequence is that of a cancer specific gene (CSG) designated pro108. The CSG was identified using a suppression subtractive hybridization method. It is a diagnostic marker for prostate cancer. In 11 different healthy tissues examined, pro108 mRNA expression was highest in ovary and uterus. Expression in healthy prostate was low. However, overexpression of CSG pro108 was observed in 13 of 13 prostate cancer tissues examined. CSGs comprising pro108, its encoded polypeptide (see AAY79561), or an antibody that specifically binds CSG, are used in claimed methods for the diagnosis, detection, staging, monitoring, imaging and treatment of prostate cancer. The new methods provide earlier diagnosis for the presence and metastasis of prostate cancer, and can be used to determine if a cancer has metastasized, or to monitor the progress or stage of the disease when it has not metastasized.

Sequence 1840 BP: 305 A: 628 C: 582 G: 325 T: 0 other:

Query Match	100.0%	Score 1840;	DB 21;	Length 1840;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1840;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	CTCAGCCGGGTGCGGCAGCACTGCCAGGGGAAGAGGTGATCGACCCCGGGAAGGTC	60		
Db 1	CTCGAGCCGGGTGCGGCAGCACTGCCAGGGGAAGAGGTGATCGACCCCGGGAAGGTC	60		
Qy 61	GCTGGGCAGGCGAGTTGGGAAAGCGGAGCCCGCCGCCCGCCCGAGCCCTTCTCCTC	120		
Db 61	GCTGGGCAGGCGAGTTGGGAAAGCGGAGCCCGCCGCCCGCCCGAGCCCTTCTCCTC	120		
Qy 121	CTTTCTCCACGTCCTATCTGCCTCTCCGTGGAGGCCAGGCGTGCAGATCGAAGACAG	180		
Db 121	CTTTCTCCACGTCCTATCTGCCTCTCCGTGGAGGCCAGGCGTGCAGATCGAAGACAG	180		
Qy 181	GAGGAACATGGAGCCCTCATTTGGCCGCCCGGGCGCGGCTCGGGCTTAAATAGGAGTCT	240		
Db 181	GAGGAACATGGAGCCCTCATTTGGCCGCCCGGGCGCGGCTCGGGCTTAAATAGGAGTCT	240		
Qy 241	CGGGCTCTGGGCTGGGACCCGACCGCTGCCGCGCGCTCCCGCTCTCTCTGCGGGTGAT	300		
Db 241	CGGGCTCTGGGCTGGGACCCGACCGCTGCCGCGCGCTCCCGCTCTCTCTGCGGGTGAT	300		
Qy 301	GGAAACCCCGAGCCGGCGCGGCGGCGCTGGGCAAGGCCCTCTGCGCTCTCTCTGGCCAC	360		
Db 301	GGAAACCCCGAGCCGGCGCGGCGGCGCTGGGCAAGGCCCTCTGCGCTCTCTCTGGCCAC	360		
Qy 361	TCTCGGCGCGCGGCCACGCTCTTTGGGGGAGAGTCATCTGTTCCGCAGAGACCCCGGC	420		
Db 361	TCTCGGCGCGCGGCCACGCTCTTTGGGGGAGAGTCATCTGTTCCGCAGAGACCCCGGC	420		
Qy 421	CAAAATACAGCATCACTTTACGGGCAAGTGGACGACAGCGGCTTCCCAAGCAGTACCC	480		
Db 421	CAAAATACAGCATCACTTTACGGGCAAGTGGACGACAGCGGCTTCCCAAGCAGTACCC	480		
Qy 481	CCTGTTCCGCGCCCTTGGCGAGTGTCTTCGTGCTGGGGCGCGCATAGTCCGACTA	540		
Db 481	CCTGTTCCGCGCCCTTGGCGAGTGTCTTCGTGCTGGGGCGCGCATAGTCCGACTA	540		
Qy 541	CAGCATGTGGAGGAACACAGTACGTCTAGTAACGGGCTGCGCGACTTTGCGGAGCGGG	600		
Db 541	CAGCATGTGGAGGAACACAGTACGTCTAGTAACGGGCTGCGCGACTTTGCGGAGCGGG	600		
Qy 601	CGAGSCCTGGGCGTGTATGAAGGAGATCGAGCGCGGGGAGGCGCTGCAGAGGCTGCA	660		
Db 601	CGAGSCCTGGGCGTGTATGAAGGAGATCGAGCGCGGGGAGGCGCTGCAGAGGCTGCA	660		

QY 1741 GAAGTGAATAAATGGGCGGTTTCGGAAGCGTCAGTGTTCATGTATGATGATCTCTG 1800
|||||
Db 1741 GAAGTGAATAAATGGGCGGTTTCGGAAGCGTCAGTGTTCATGTATGATGATCTCTG 1800
QY 1801 CGTTTGAATAAAGACATCTCTGTTGCTCAAAAAA 1840
|||||
Db 1801 CGTTTGAATAAAGACATCTCTGTTGCTCAAAAAA 1840
RESULT 2
AA158028
ID AA158028 standard; cDNA; 1991 BP.
XX
AC AA158028;
DT
XX 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 231.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR P-PSDB; AAM38872.
DR
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
PS Claim 1; SEQ ID NO 231; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC
CC Note: The sequence data for this patent did not form part of the printed

CC specification.
XX
SQ Sequence 1991 BP; 339 A; 669 C; 620 G; 350 T; 13 other;
Query Match 99.3%; Score 1827.6; DB 22; Length 1991;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 7 CCGGGGTGCGGCAGCACTGCCAGGGAAGAGGGGTGATCCGACCCGGGGAAGGTGCGTGGG 66
|||||
Db 97 CCGGGGTGCGGCAGCACTGCCAGGGAAGAGGGGTGATCCGACCCGGGGAAGGTGCGTGGG 156
QY 67 CAGGGCGAGTTGGAAAGCGGACGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 126
|||||
Db 157 CAGGGCGAGTTGGAAAGCGGACGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 216
QY 127 CCCACGTCTATCTGCTCTGCTGAGGCGAGGCGGCTGCGAGCATCGAAGACAGGAGAA 186
|||||
Db 217 CCCACGTCTATCTGCTCTGCTGAGGCGAGGCGGCTGCGAGCATCGAAGACAGGAGAA 276
QY 187 CTGGAGCCTCATTTGGCGGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 246
Db 277 CTGGAGCCTCATTTGGCGGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 336
QY 247 CTGGCTGGGACCGACCGCTGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 306
Db 337 CTGGCTGGGACCGACCGCTGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 396
QY 307 CCCACGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 366
Db 397 CCCACGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 456
QY 367 CGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 426
Db 457 CGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 516
QY 427 CAGCATACCTTTCACGGGCAAGTGGAGCCAGACGGGCTTCCCAAGCAGTACCCCTGTT 486
Db 517 CAGCATACCTTTCACGGGCAAGTGGAGCCAGACGGGCTTCCCAAGCAGTACCCCTGTT 576
QY 487 CCGCCGCCCTGCGCAGTGTCTTTCGCTGCTGGGGCGGCGCATAGTCTCCGACTACAGCAT 546
Db 577 CCGCCGCCCTGCGCAGTGTCTTTCGCTGCTGGGGCGGCGCATAGTCTCCGACTACAGCAT 636
QY 547 GTGGAGGAACCACTAGTACGTACAGTGAAGCGCTGCGGCTGCGGAGCGCGCGGCGGCG 606
Db 637 GTGGAGGAACCACTAGTACGTACAGTGAAGCGCTGCGGCTGCGGAGCGCGCGGCGGCG 696
QY 607 CTGGCGCTGATGAAGGAGATCGAGCGCGGGGAGGCGCTGCAGAGCGTGCACGAGGT 666
Db 697 CTGGCGCTGATGAAGGAGATCGAGCGCGGGGAGGCGCTGCAGAGCGTGCACGCGGT 756
QY 667 GTTTTCGGGCGCGCGTCCCGAGCGCACCGGGCGAGACGTGCGGCGAGGTGAGGTGCA 726
Db 757 GTTTTCGGGCGCGCGTCCCGAGCGCACCGGGCGAGACGTGCGGCGAGGTGAGGTGCA 816
QY 727 GCGCAGGCACTCGCTGCTCTGTTGTGGTGGCGCATCTGCCAGCCCCGACTGTTCTGT 786
Db 817 GCGCAGGCACTCGCTGCTCTGTTGTGGTGGCGCATCTGCCAGCCCCGACTGTTCTGT 876
QY 787 GGGCGTGGACAGCTGGACCTGTGCCAGGGGACCGTTGGCGGGAACAGCGCGCGCTGGA 846
Db 877 GGGCGTGGACAGCTGGACCTGTGCCAGGGGACCGTTGGCGGGAACAGCGCGCGCTGGA 936
QY 847 CTGTGTACCCCTAGCAGCGCGGGGACGGACAGCGGCTTACCTTCTCTCCCGAACCTTCG 906
Db 937 CTGTGTACCCCTAGCAGCGCGGGGACGGACAGCGGCTTACCTTCTCTCCCGAACCTTCG 996
QY 907 CACCATCCCGCAGGACACGGTGTACCCAGAGATAAGTCTCTCTCCAGCCACCGCGGCAA 966
Db 997 CACCATCCCGCAGGACACGGTGTACCCAGAGATAAGTCTCTCTCTCCAGCCACCGCGGCAA 1056
QY 967 CTCCTTCTACTACCCGCGGCTGAAGGCCCTGCTCCCATCGCCAGGGGTGACACTGGTGGC 1026

Db 1057 CTCTCTTACTACCGCGGTGAAGCCCTGCTCCCTCCATCCGAGGCTGACACTGTGCG 1116
QY 1027 GCTGCGACAGAGCCAGGCGCTTCATCCCTCCCGCCAGTCTGCCAGCAGGAGCAA 1086
Db 1117 GCTGCGACAGAGCCAGGCGCTTCATCCCTCCCGCCAGTCTGCCAGCAGGAGCAA 1176
QY 1087 TGAGATTGTAGACAGCGCTTCAGTCCAGAAACCGCGCTGGAGTGGAGGTCTCCCTGTG 1146
Db 1177 TGAGATTGTAGACAGCGCTTCAGTCCAGAAACCGCGCTGGAGTGGAGGTCTCCCTGTG 1236
QY 1147 GTGCTCTGGGACTGTGCGGAGGCGCACTGTGGAGGCTCGGACCAAGAGGAGTCTG 1206
Db 1237 GTGCTCTGGGACTGTGCGGAGGCGCACTGTGGAGGCTCGGACCAAGAGGAGTCTG 1296
QY 1207 CTAGCTCCGGGTCCAGCCGCCAACAGGAGCGCTCGCCCGAGCTCGAAGAAAGGCG 1266
Db 1297 CTAGCTCCGGGTCCAGCCGCCAACAGGAGCGCTCGCCCGAGCTCGAAGAAAGGCG 1356
QY 1267 TGAGTGGGTCCCTGATAACTGCTCTAAGACACAGAGCCCGCAGCCCTGGGCCCCCG 1326
Db 1357 TGAGTGGGTCCCTGATAACTGCTCTAAGACACAGAGCCCGCAGCCCTGGGCCCCCG 1416
QY 1327 GAGCCATGGGTGTGGGGGCTCTGTGCAAGGCTCATGCTGCAAGGCGCGGAGGACAG 1386
Db 1417 GAGCCATGGGTGTGGGGGCTCTGTGCAAGGCTCATGCTGCAAGGCGCGGAGGACAG 1476
QY 1387 GGGTGTGGGTCTGCTGACCGCGGTGAGCGCGCGCGGCGGACCATCTCTCACTGAAGG 1446
Db 1477 GGGTGTGGGTCTGCTGACCGCGGTGAGCGCGCGCGGCGGACCATCTCTCACTGAAGG 1536
QY 1447 CCCTCTGGTGGCGGCGGCGGATGGGAAACAGCCCTCTCTCTTCCCAACCTTGCTTCT 1506
Db 1537 CCCTCTGGTGGCGGCGGCGGATGGGAAACAGCCCTCTCTCTTCCCAACCTTGCTTCT 1596
QY 1507 TAGGGCCCCCGTGTCCGCTCTGCTCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1566
Db 1597 TAGGGCCCCCGTGTCCGCTCTGCTCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1656
QY 1567 CAAGGTCTCAGCTACTTAATATATGCTCTCTTATAAGTTATTGCTCTCCAGGAGTG 1626
Db 1657 CAAGGTCTCAGCTACTTAATATATGCTCTCTTATAAGTTATTGCTCTCCAGGAGTG 1716
QY 1627 TCCTTCATCTCCAGGGGCTGCTCCAGCGTGTGCGATACCTCAGACCTGCTGCTC 1686
Db 1717 TCCTTCATCTCCAGGGGCTGCTCCAGCGTGTGCGATACCTCAGACCTGCTGCTC 1776
QY 1687 TAGGCTGTGTAGCCCACTCTCCGAGGCGCATCCAGCGGGGCGCACTTGAGAAGTG 1746
Db 1777 TAGGCTGTGTAGCCCACTCTCCGAGGCGCATCCAGCGGGGCGCACTTGAGAAGTG 1836
QY 1747 AATAAATGGGCGGTTCGGAAGCGTCAGTGTTCATGTTTCCATGTTATGATCTCTCGCTTG 1806
Db 1837 AATAAATGGGCGGTTCGGAAGCGTCAGTGTTCATGTTTCCATGTTATGATCTCTCGCTTG 1896
QY 1807 AATAAGACTATCTCTGTTGCTCAAAAAA 1840
Db 1897 AATAAGACTATCTCTGTTGCTCAAAAAA 1930

RESULT 3

AAI59814/c

ID AAI59814 standard; cdna; 1926 BP.

AC AAI59814;

XX AAI59814;

XX 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 3803.

DE Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW

amytrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
Homo sapiens.
WO200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US34263.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;
WPI: 2001-442253/47.
P-PSDB; AAM40858.
Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
Claim 1; SEQ ID NO 3803; 10078pp; English.
The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nontropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed specification.
SQ Sequence 1926 BP; 337 A; 604 C; 645 G; 334 T; 6 other;
Query Match 99.2%; Score 1825.6; DB 22; Length 1926;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1828; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 9 GGGGTGCGGCGACACTGCCAGGGAAGAGGGTGATCCGACCCGGGGAAGGTCTCGTGGCA 68
Db 1835 GGGGTGCGGCGACACTGCCAGGGAAGAGGGTGATCCGACCCGGGGAAGGTCTCGTGGCA 1776
QY 69 GGGCGAGTTGGGAAAGCGCAGCCCGCCGCGCCGCGAGCCCTTCTCTCTCTCTCTCC 128
Db 1775 GGGCGAGTTGGGAAAGCGCAGCCCGCCGCGCCGCGAGCCCTTCTCTCTCTCTCTCC 1716
QY 129 CAGTCTCTATCTGCTCTCTGCTGGAGCCGCGCTGCGAGTCGAGAGGAGGAAC 188
Db 1715 CAGTCTCTATCTGCTCTCTGCTGGAGCCGCGCTGCGAGTCGAGAGGAGGAAC 1656
QY 189 GGAGCCCTCATTTGGCCGGCGCGGCGCGCTCGGGCTTAATAGAGGCTCGGGCTCT 248
Db 1655 GGAGCCCTCATTTGGCCGGCGCGGCGCGCTCGGGCTTAATAGAGGCTCGGGCTCT 1596

PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 23-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 27-APR-1998; 98US-0082796.
PR 28-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 18-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 30-JUL-1998; 98US-0094651.
PR

PR 11-SEP-1998; 98US-0100038.
XX (GETH) GENENTECH INC.
XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX WPI; 1999-551358/46.
DR P-PSDB; AAY41721.
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
PS Claim 2; Fig 86; 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AA33891 to
CC AA234338, and AAY41685 to AAY41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
XX invention.
XX
SQ Sequence 1846 BP; 331 A; 622 C; 571 G; 322 T; 0 other;

Query Match 98.5%; Score 1812.8; DB 20; Length 1846;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1814; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 GCCAGGGGAGAGGCTGATCCGACCGCGGGAAGGTGCGTGGCAGGCGCAGTTGGGAAG 84
DB 1 GCCAGGGGAGAGGCTGATCCGACCGCGGGAAGGTGCGTGGCAGGCGCAGTTGGGAAG 60

QY 85 CGCGAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 144
DB 61 CGCGAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 120

QY 145 CTCGCTGGAGGCGCAGGCGCTGCGAGCATCGAACAGAGGAACTGGAGCCTCATTTGGCG 204
DB 121 CTCGCTGGAGGCGCAGGCGCTGCGAGCATCGAACAGAGGAACTGGAGCCTCATTTGGCG 180

QY 205 GCCGGGCGCGCGCCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCG 264
DB 181 GCCGGGCGCGCGCCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCG 240

QY 265 CTCGCGCGCGCGCTCCCGCTGCTCTGCGGGGTGATGGAACCCCGCCCGCCCGCCCG 324
DB 241 CTCGCGCGCGCGCTCCCGCTGCTCTGCGGGGTGATGGAACCCCGCCCGCCCGCCCG 300

QY 325 CTTGGCAAGGCGCTCTGCGCTCTCTCTGCGCACTCTGCGGCGCGCGCGCGCGCGCTCT 384
DB 301 CTTGGCAAGGCGCTCTGCGCTCTCTCTGCGCACTCTGCGGCGCGCGCGCGCGCTCT 360

QY 385 TGGGGGAGAGTCCACTGTTCCGCCAGAGCCCGCGCCAAATACAGCATCACCTTACCGG 444
DB 361 TGGGGGAGAGTCCACTGTTCCGCCAGAGCCCGCGCCAAATACAGCATCACCTTACCGG 420

QY 445 CAAGTGGAGCCAGACGGCTTCCCCAAGCAGTACCCCTCTTCCCGCCCGCCCTGCGCAGTG 504
DB 421 CAAGTGGAGCCAGACGGCTTCCCCAAGCAGTACCCCTCTTCCCGCCCGCCCTGCGCAGTG 480

QY 505 GTCTTCGCTGCTGGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGAAAGAACAGTA 564
DB 481 GTCTTCGCTGCTGGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGAAAGAACAGTA 540

QY 565 CGTCACTAACGGGCTGCGGACTTTTCGGAGGCGCGGAGGCTGGGCGCTGATGAAGA 624
DB 541 CGTCACTAACGGGCTGCGGACTTTTCGGAGGCGCGGAGGCTGGGCGCTGATGAAGA 600

The present invention describes secreted and transmembrane polypeptides and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AA33891 to AA234338, and AAY41685 to AAY41774 represent polynucleotide and polypeptide sequence given in the exemplification of the present invention.

QY 625 GATGAGGCGGGGGAGCGCTGCAGAGCGTGACAGAGGTGTTTTCGGCGCGCGCTT 684
|||||
Db 601 GATCAGGCGGGGGAGCGCTGCAGAGCGTGACAGAGGTGTTTTCGGCGCGCGCT 660
QY 685 CCCAGCGCACCGGCGACAGCTCGCGGAGCTGAGGTGACAGCGACATCGCTGT 744
|||||
Db 661 CCCAGCGCACCGGCGACAGCTCGCGGAGCTGAGGTGACAGCGACATCGCTGT 720
QY 745 CTCGTTGTGTGGCTGCGCATGTCGCCAGCCCGACCTGTTCTGTTGGCGTGACAGCTGGA 804
|||||
Db 721 CTCGTTGTGTGGCTGCGCATGTCGCCAGCCCGACCTGTTCTGTTGGCGTGACAGCTGGA 780
QY 805 CTTGTCCGACGGGAGCGCTTGGCGGACACAGCGGCGCTGGACCTGTACCCCTACGACGC 864
|||||
Db 781 CTTGTCCGACGGGAGCGCTTGGCGGACACAGCGGCGCTGGACCTGTACCCCTACGACGC 840
QY 865 CGGGACGGACAGCGGTTTCACTTCTCTCCCGCCCACTTTCGGCCACATCCCGCAGGACAC 924
|||||
Db 841 CGGGACGGACAGCGGTTTCACTTCTCTCCCGCCCACTTTCGGCCACATCCCGCAGGACAC 900
QY 925 GGTGACCGAGATAAGCTCTCTCTCCAGCCACCGCCCACTCTTCTACTACCCGG 984
|||||
Db 901 GGTGACCGAGATAAGCTCTCTCTCCAGCCACCGCCCACTCTTCTACTACCCGG 960
QY 985 GCTGAAGGCCCTCGCTCCCATCGCAGGCTGACACTGTGTGCGGCTCCGACAGAGCCCCAG 1044
|||||
Db 961 GCTGAAGGCCCTCGCTCCCATCGCAGGCTGACACTGTGTGCGGCTCCGACAGAGCCCCAG 1020
QY 1045 GGCCTTCATCCTCCCGCCCCAGTCTGCCAGAGGAGCAATGATGTTAGACAGCGC 1104
|||||
Db 1021 GGCCTTCATCCTCCCGCCCCAGTCTGCCAGAGGAGCAATGATGTTAGACAGCGC 1080
QY 1105 CTCAGTCTCAGAAAGCCGCTGACTCGAGGCTCTCCCTGTGTGTCTGTGGGACTGTG 1164
|||||
Db 1081 CTCAGTCTCAGAAAGCCGCTGACTCGAGGCTCTCCCTGTGTGTCTGTGGGACTGTG 1140
QY 1165 CGGAGGCCACTGTGGGAGGCTCGGACCAAGACGAGGACTCGCTACGTCGGGTCCAGCC 1224
|||||
Db 1141 CGGAGGCCACTGTGGGAGGCTCGGACCAAGACGAGGACTCGCTACGTCGGGTCCAGCC 1200
QY 1225 CGCCAAACAGGAGGCCCTGCCCGAGCTCGAAGAAGAGGCTGAGTGGCTCCCTGATAA 1284
|||||
Db 1201 CGCCAAACAGGAGGCCCTGCCCGAGCTCGAAGAAGAGGCTGAGTGGCTCCCTGATAA 1260
QY 1285 CTGCGTCTAAGACACAGAGCCCGACGCCCTGSGGCCCGCCGAGCCATGGGTGTCGG 1344
|||||
Db 1261 CTGCGTCTAAGACACAGAGCCCGACGCCCTGSGGCCCGCCGAGCCATGGGTGTCGG 1320
QY 1345 GGCTCCTGTGCAGGCTCATGCTGCAGCGCGCGAGGGCACAGGGGGTTTTCGGGCTGCTCC 1404
|||||
Db 1321 GGCTCCTGTGCAGGCTCATGCTGCAGCGCGCGAGGGCACAGGGGGTTTTCGGGCTGCTCC 1380
QY 1405 TGACCGCGGTGAGCGCGCGCGGACCATCTCTGACTGAAGGCCCTCTGGTGCGCGGCAC 1464
|||||
Db 1381 TGACCGCGGTGAGCGCGCGCGGACCATCTCTGACTGAAGGCCCTCTGGTGCGCGGCAC 1440
QY 1465 GGGCATTGGGAACAGCCCTCTCTTCCCAACCTTGCTTCTAGGGCCCCCGTCC 1524
|||||
Db 1441 GGGCATTGGGAACAGCCCTCTCTTCCCAACCTTGCTTCTAGGGCCCCCGTCC 1500
QY 1525 GTCTGCTCTCAGCCCTCTCTCTCTGAGGATAAAGTCACTCCCAAGGCTCCAGCTACTCT 1584
|||||
Db 1501 GTCTGCTCTCAGCCCTCTCTCTCTGAGGATAAAGTCACTCCCAAGGCTCCAGCTACTCT 1560
QY 1585 AAATTATGTCTCTTATAGTTATTGCTGCTCCAGGAGATTGCTCTTATCGTCCAGGGG 1644
|||||
Db 1561 AAATTATGTCTCTTATAGTTATTGCTGCTCCAGGAGATTGCTCTTATCGTCCAGGGG 1620
QY 1645 CTTGGCTCCACGTTGTCATACCTCAGACCTGGTGCTCTAGCTGTGCTGAGCCCA 1704
|||||
Db 1621 CTTGGCTCCACGTTGTTGATACCTCAGACCTGGTGCTCTAGCTGTGCTGAGCCCA 1680
QY 1705 CTCTCCGAGGGCGCATCCAAAGGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTTC 1764

Db 1681 CTCCTCCGAGGCGCATCCAAAGCGGGCCACTTGAGAAGTGAATAAATGGGCGGTTTC 1740
QY 1765 GGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGCGGTTTGAATAAAGACTATCTCTGT 1824
|||||
Db 1741 GGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGCGGTTTGAATAAAGACTATCTCTGT 1800
QY 1825 TGCTCAAAAAA 1840
Db 1801 TGCTCAAAAAA 1816
RESULT 5
AAC78507
ID AAC78507 standard; cDNA; 1847 BP.
XX
AC AAC78507;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO866 (UNQ435) nucleotide sequence SEQ ID NO:235.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
XX expressed sequence tag; detection; cancer; ss.
OS Homo sapiens.
XX
PN WO200053756-A2.
XX
PD 14-SEP-2000.
XX
PF 18-FEB-2000; 2000WO-US04341.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kijavini IJ, Kuo SS, Napier WA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX
DR WPI; 2000-611443/58.
DR P-PSDB; AAB44277.
XX
PT Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities .
XX
PS Claim 2; Fig 86; 636pp; English.
XX
CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting

CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX
SQ Sequence 1847 BP; 332 A; 622 C; 571 G; 322 T; 0 other;

Query Match 98.5%; Score 1812.8; DB 21; Length 1847;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1814; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	25	GCCAGGGGAGAGGGTATCCGACCCGGGGAAGGTCGCTGGGCAGGCGGAGTTGGGAAG	84
DB	1	GCCAGGGGAGAGGGTATCCGACCCGGGGAAGGTCGCTGGGCAGGCGGAGTTGGGAAG	60
QY	85	CGGACGCCCCCGCGCCCGCGAGCCCTTCTCCTCTTCTCCACGTCCTATCTGCCT	144
DB	61	CGGACGCCCCCGCGCCCGCGAGCCCTTCTCCTCTTCTCCACGTCCTATCTGCCT	120
QY	145	CTCGCTGGAGGCCAGCGCTGACGATCGAAGACAGAGGAACTGAGCCTCATTTGGCCG	204
DB	121	CTCGCTGGAGGCCAGCGCTGACGATCGAAGACAGAGGAACTGAGCCTCATTTGGCCG	180
QY	205	GCCCGGGGCGCGCCCTCGGGCTTAATAAGAGCTCCGGCTCTGGCTGGGACCCGACCG	264
DB	181	GCCCGGGGCGCGCCCTCGGGCTTAATAAGAGCTCCGGCTCTGGCTGGGACCCGACCG	240
QY	265	CTCGCGGCGCGCTCCGCTGCTCTGCGGGTGATGGAACCCAGCCCGCGCCGCG	324
DB	241	CTCGCGGCGCGCTCCGCTGCTCTGCGGGTGATGGAACCCAGCCCGCGCCGCGCG	300
QY	325	CTGCGCAAGGCCCTCTGCGCTCTCTCTGCGCACTCTCGGGCGCGCGCGCAGCTCT	384
DB	301	CTGCGCAAGGCCCTCTGCGCTCTCTCTGCGCACTCTCGGGCGCGCGCGCAGCTCT	360
QY	385	TGGGGGAGAGTCCATCTGTTCCGCCAGAGCCCGGCAAAATACAGCATCACCTTACGGG	444
DB	361	TGGGGGAGAGTCCATCTGTTCCGCCAGAGCCCGGCAAAATACAGCATCACCTTACGGG	420
QY	445	CAAGTGGAGCAGACGGCTTCCCAAGCAGTACCCCTCTTCCGCCCGCCCTGCGCAGTG	504
DB	421	CAAGTGGAGCAGACGGCTTCCCAAGCAGTACCCCTCTTCCGCCCGCCCTGCGCAGTG	480
QY	505	GTCTTCGCTGCTGGGGCGCGCATAGCTCCGACTACAGCATGTGAGGAAGAACAGTA	564
DB	481	GTCTTCGCTGCTGGGGCGCGCATAGCTCCGACTACAGCATGTGAGGAAGAACAGTA	540
QY	565	CGTCAGTAACGGGCTCGCGGACTTTTCGGAGCGCGCGGAGGCTGGCGCTGATGAAGA	624
DB	541	CGTCAGTAACGGGCTCGCGGACTTTTCGGAGCGCGCGGAGGCTGGCGCTGATGAAGA	600
QY	625	GATCGAGGCGGGGGAGGCGCTGCAGAGCGTGACAGAGGTGTTTTCGGCGCCCGCCGT	684
DB	601	GATCGAGGCGGGGGAGGCGCTGCAGAGCGTGACAGAGGTGTTTTCGGCGCCCGCCGT	660
QY	685	CCCCAGCGCACCGGCGACGCTCGGCGAGCTGGAGGTGACGCGACGACCTCGTGTGT	744
DB	661	CCCCAGCGCACCGGCGACGCTCGGCGAGCTGGAGGTGACGCGACGACCTCGTGTGT	720
QY	745	CTCGTTTGTGTCGCATCGTGCCAGCCCGACTGGTTCTGTTGGGCGTGACAGCCTGGA	804
DB	721	CTCGTTTGTGTCGCATCGTGCCAGCCCGACTGGTTCTGTTGGGCGTGACAGCCTGGA	780
QY	805	CCTGTCCGAGGGGACCGTTGGGGGAAACAGGGGCGCTGGACCTGTACCCCTACGACGC	864
DB	781	CCTGTCCGAGGGGACCGTTGGGGGAAACAGGGGCGCTGGACCTGTACCCCTACGACGC	840
QY	865	CGGGACGGACAGCGGCTTCACTTCTCTCCCGCAACTTCGCCACCATCCCGCAGACAC	924
DB	841	CGGGACGGACAGCGGCTTCACTTCTCTCCCGCAACTTCGCCACCATCCCGCAGACAC	900

QY	925	GGTGACCGAGATAACGTCTCTCTCTCCAGCCACCGGCAACTCTTCTACTACCCGCG	984
DB	901	GGTGACCGAGATAACGTCTCTCTCTCCAGCCACCGGCAACTCTTCTACTACCCGCG	960
QY	985	GCTGAAGGCCCTGCTCTCCATGCCAGGTGACACTGTGTGGCTGCGACAGAGCCCCAG	1044
DB	961	GCTGAAGGCCCTGCTCTCCATGCCAGGTGACACTGTGTGGCTGCGACAGAGCCCCAG	1020
QY	1045	GGCCTTCATCCCTCCGCCCCAGTCTGCCCCAGGAGCAATGAGATTGTAGACAGCGC	1104
DB	1021	GGCCTTCATCCCTCCGCCCCAGTCTGCCCCAGGAGCAATGAGATTGTAGACAGCGC	1080
QY	1105	CTCAGTTCCAGAAACCGCTGGACTGCGAGGTCTCCCTGTGTCTCTCTGGGACTGTG	1164
DB	1081	CTCAGTTCCAGAAACCGCTGGACTGCGAGGTCTCCCTGTGTCTCTCTGGGACTGTG	1140
QY	1165	CGGAGGCCACTGTGGAGGCTCGGGACCAAGAGAGCTCGCTACGTCGGGTCCAGCC	1224
DB	1141	CGGAGGCCACTGTGGAGGCTCGGGACCAAGAGAGCTCGCTACGTCGGGTCCAGCC	1200
QY	1225	CGCCAAACGGGAGGCCCTGCCCCAGCTCGAAGAGAGGCTGAGTGGTCCCTGATAA	1284
DB	1201	CGCCAAACGGGAGGCCCTGCCCCAGCTCGAAGAGAGGCTGAGTGGTCCCTGATAA	1260
QY	1285	CTGCGTCTAAGACACAGAGCCCGCGAGCCCTGGGGCCCCCGGAGCATGGGTGTGCGG	1344
DB	1261	CTGCGTCTAAGACACAGAGCCCGCGAGCCCTGGGGCCCCCGGAGCATGGGTGTGCGG	1320
QY	1345	GGTCTCTGTGAGGCTCATGTGCGAGCGCGGAGGCGACAGGGGGTTTCGGCTGCTCC	1404
DB	1321	GGTCTCTGTGAGGCTCATGTGCGAGCGCGGAGGCGACAGGGGGTTTCGGCTGCTCC	1380
QY	1405	TGACGCGGTGAGGCGCGCGGAGCATCTGCACTGAAGGGCCCTCTGCTGGCGGCGAC	1464
DB	1381	TGACGCGGTGAGGCGCGCGGAGCATCTGCACTGAAGGGCCCTCTGCTGGCGGCGAC	1440
QY	1465	GGGCATTGGGAACAGCCCTCCTTTCCCAACCTTTCCTTAGGGGCCCCCGTGTCCC	1524
DB	1441	GGGCATTGGGAACAGCCCTCCTTTCCCAACCTTTCCTTAGGGGCCCCCGTGTCCC	1500
QY	1525	GTCTGCTCTCAGGCTCCTCTCTGCGAGGATAAAGTCATCCCAAGGCTCCAGCTACTCT	1584
DB	1501	GTCTGCTCTCAGGCTCCTCTCTGCGAGGATAAAGTCATCCCAAGGCTCCAGCTACTCT	1560
QY	1585	AAATTATGTCCTTATTAAGTTATTCCTGCTCCAGAGATTGCTTCATCGTCCAGGG	1644
DB	1561	AAATTATGTCCTTATTAAGTTATTCCTGCTCCAGAGATTGCTTCATCGTCCAGGG	1620
QY	1645	CCTGGCTCCAGCTGGTTGCAGATACCTCAGACCTGGTGTCTTAGGCTGTGTGAGCCCA	1704
DB	1621	CCTGGCTCCAGCTGGTTGCAGATACCTCAGACCTGGTGTCTTAGGCTGTGTGAGCCCA	1680
QY	1705	CTCTCCGAGGCGCATCCAAAGCGGGGCGACTTGAGAAGTGAATAAATGGGCGGTTTC	1764
DB	1681	CTCTCCGAGGCGCATCCAAAGCGGGGCGACTTGAGAAGTGAATAAATGGGCGGTTTC	1740
QY	1765	GGAAGGTCAGTGTTCCTATGATTTATGATCTCTGCGTTTGAATAAAGACTATCTCTGT	1824
DB	1741	GGAAGGTCAGTGTTCCTATGATTTATGATCTCTGCGTTTGAATAAAGACTATCTCTGT	1800
QY	1825	TGCTCAAAAAAANA 1840	
DB	1801	TGCTCAAAAAAANA 1816	

RESULT 6
AAC58630
ID AAC58630 standard; cDNA; 1847 BP.
XX
AC AAC58630;
XX
DT 29-JAN-2001 (first entry)

XX DE Human PRO866 protein UNQ435 encoding cDNA SEQ ID NO:257.

XX KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;

XX KW dermatological; antiarthritic; antirheumatic; immunosuppressive;

XX KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;

XX KW antianemic; hepatotropic; virucide; antipsoriatic; antiallergic;

XX KW antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;

XX KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;

XX KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;

XX KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;

XX KW autoimmune thrombocytopaenia; immune-mediated renal disease;

XX KW demyelinating disease; hepatobiliary disease; Whipple's disease;

XX KW inflammatory bowel disease; gluten-sensitive enteropathy;

XX KW autoimmune disease; immune-mediated skin disease; allergic disease;

XX KW immunological disease; transplantation associated disease;

XX KW graft rejection; graft-versus-host-disease; ss.

OS Homo sapiens.

XX WO2000053758-A2.

XX PD 14-SEP-2000.

XX 02-MAR-2000; 2000WO-US05841.

XX 08-MAR-1999; 99WO-US05028.

XX 10-MAR-1999; 99US-0123618.

XX 12-MAR-1999; 99US-0123957.

XX 23-MAR-1999; 99US-0125775.

XX 12-APR-1999; 99US-0128848.

XX 20-APR-1999; 99WO-US08615.

XX 28-APR-1999; 99US-0131445.

XX 04-MAY-1999; 99US-0132371.

XX 14-MAY-1999; 99US-0134287.

XX 02-JUN-1999; 99WO-US12252.

XX 23-JUN-1999; 99US-0141037.

XX 20-JUL-1999; 99US-0144758.

XX 26-JUL-1999; 99US-0145698.

XX 28-JUL-1999; 99US-0146222.

XX 01-SEP-1999; 99WO-US20111.

XX 08-SEP-1999; 99WO-US20594.

XX 13-SEP-1999; 99WO-US20944.

XX 15-SEP-1999; 99WO-US21090.

XX 15-SEP-1999; 99WO-US21547.

XX 05-OCT-1999; 99WO-US23089.

XX 29-OCT-1999; 99US-0162506.

XX 29-NOV-1999; 99WO-US28214.

XX 30-NOV-1999; 99WO-US28313.

XX 30-NOV-1999; 99WO-US28409.

XX 01-DEC-1999; 99WO-US28301.

XX 01-DEC-1999; 99WO-US28634.

XX 02-DEC-1999; 99WO-US28511.

XX 02-DEC-1999; 99WO-US28564.

XX 12-DEC-1999; 99WO-US28565.

XX 16-DEC-1999; 99WO-US30095.

XX 20-DEC-1999; 99WO-US30999.

XX 30-DEC-1999; 99WO-US31274.

XX 05-JAN-2000; 2000WO-US00219.

XX 06-JAN-2000; 2000WO-US00277.

XX 06-JAN-2000; 2000WO-US00376.

XX 11-FEB-2000; 2000WO-US03565.

XX 18-FEB-2000; 2000WO-US04341.

XX 18-FEB-2000; 2000WO-US04342.

XX 22-FEB-2000; 2000WO-US04414.

XX (GETH) GENENTECH INC.

XX AShenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;

XX Kabakoff RC, Lu Y, Pennica D, Shelton DL, Smith V;

XX Stewart TA, Tumas D, Watanabe CK, Wood WL, Yan M;

XX WPI; 2000-572271/53.

DR P-PSDB; AAB33465.

XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of

PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid

PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -

XX Claim 23; Fig 103; 309pp; English.

XX The present invention describes sixty four human PRO proteins which can

CC be used in the treatment of immune related diseases. The human PRO

CC proteins, anti-PRO antibodies, agonists and antagonists are useful for

CC treating and diagnosing immune related disorders. The disorders are

CC selected from systemic lupus erythematosus, rheumatoid arthritis,

CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,

CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's

CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic

CC anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,

CC immune-mediated renal disease, demyelinating diseases of the central

CC and peripheral nervous systems, hepatobiliary diseases, inflammatory

CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,

CC autoimmune or immune-mediated skin diseases, allergic diseases,

CC immunological diseases of the lung, and transplantation associated

CC diseases including graft rejection and graft-versus-host-disease.

CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used

CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and

CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein

CC sequences given in the exemplification of the present invention.

XX

SQ Sequence 1847 BP; 332 A; 622 C; 571 G; 322 T; 0 other;

Query Match 98.5%; Score 1812.8; DB 21; Length 1847;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1814; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 GCCAGGGGAAGAGGTGATCCGACCCGGGGAAGTTCGTGGCAGGGCGAGTGGGAAG 84

DB 1 GCCAGGGGAAGAGGTGATCCGACCCGGGGAAGTTCGTGGCAGGGCGAGTGGGAAG 60

QY 85 CGGCAGCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 144

DB 61 CGGCAGCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 120

QY 145 CTCGCTGGAGGCCAGCGCGCTGCAGCATCGAAGACAGAGGAGGAAGTTCGAGCTCATTTGGCGG 204

DB 121 CTCGCTGGAGGCCAGCGCGCTGCAGCATCGAAGACAGAGGAGGAAGTTCGAGCTCATTTGGCGG 180

QY 205 GCCCGGGCGCGCGCGCTCGGGCTTAAATAGAGAGCTCCGGGCTCTGGCTGGGACCGACCG 264

DB 181 GCCCGGGCGCGCGCGCTCGGGCTTAAATAGAGAGCTCCGGGCTCTGGCTGGGACCGACCG 240

QY 265 CTGCGCGCGCGCGCTCCGCTGCTCCTCGCGGTGATGGAACACCCAGCCCGCGCGCGCG 324

DB 241 CTGCGCGCGCGCGCTCCGCTGCTCCTCGCGGTGATGGAACACCCAGCCCGCGCGCGCG 300

QY 325 CCGTGGCAAGGCCCTCTGGCTCTCCTCTCGCCACTCTCGCGCGCGCGCGCGCGCGCTCT 384

DB 301 CCGTGGCAAGGCCCTCTGGCTCTCCTCTCGCCACTCTCGCGCGCGCGCGCGCGCTCT 360

QY 385 TGGGGGAGAGTCCATCTGTTCGGCAGAGCCCGGCCAAATACAGCATACACCTTCACGGG 444

DB 361 TGGGGGAGAGTCCATCTGTTCGGCAGAGCCCGGCCAAATACAGCATACACCTTCACGGG 420

QY 445 CAAAGTGGAGCCAGAGCGGCTTCCCAAGAGAGTACCCCTCTGTCGGCCCGCCCTGGCAGTG 504

DB 421 CAAAGTGGAGCCAGAGCGGCTTCCCAAGAGAGTACCCCTCTGTCGGCCCGCCCTGGCAGTG 480

QY 505 GTCTTCGCTGTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTA 564

DB 481 GTCTTCGCTGTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTA 540

QY 565 CGTCAGTAACGGGCTCGCGACTTTTGGGAGCGCGGAGGCGCTGGCGCTGATGAAGA 624

DB 541 CGTCAGTAACGGGCTCGCGACTTTTGGGAGCGCGGAGGCGCTGGCGCTGATGAAGA 600

CC A claimed method for inhibiting the growth of a tumour cell comprises
CC exposing the tumor cell to PRO179, PRO207, PRO320, PRO321, PRO221,
CC PRO224, PRO328, PRO301, PRO356, PRO362, PRO356, PRO509 or PRO866 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukaemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX
SQ Sequence 1847 BP; 332 A; 622 C; 571 G; 322 T; 0 other;

Query Match 98.5%; Score 1812.8; DB 21; Length 1847;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1814; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 GCCAGGGGAAGAGGTGATCCGACCCGCGGGAAGGTGCTGGGAGGCGAGTTGGGAAAG 84
DB 1 GCCAGGGGAAGAGGGTGATCCGACCCGCGGGAAGGTGCTGGGAGGCGAGTTGGGAAAG 60
QY 85 CGGCAGCCCCCGCCCGCCCGCGAGCCCTTCTCCTCCTTCTCCACGCTCTATCTGCCT 144
DB 61 CGGAGCCCCCGCCCGCCCGAGCCCTTCTCCTTCTCCACGCTCTATCTGCCT 120
QY 145 CTCGCTGGAGCCAGGCGCTGCAAGCATCGAAGACAGAGGAAGTGGAGCCCTATTGGCCG 204
DB 121 CTCGCTGGAGCCAGGCGCTGCAAGCATCGAAGACAGAGGAAGTGGAGCCCTATTGGCCG 180
QY 205 GCCCGGGGCGCCGCTCGGGCTTAATAGGAGCTCGGGCTTGGTGGGACCCGACCG 264
DB 181 GCCCGGGGCGCCGCTCGGGCTTAATAGGAGCTCGGGCTTGGTGGGACCCGACCG 240
QY 265 CTCGCGGCGCGCTCGCGTCTCTCGCGGTGATGGAAGAACCCAGCCGCGCCCGC 324
DB 241 CTCGCGGCGCGCTCGCGTCTCTCGCGGTGATGGAAGAACCCAGCCGCGCCCGC 300
QY 325 CTTGGGCAAGCCCTCTGCGCTCTCTCTCGGCACTCTCGGCGCGCCGCGCGAGCTCT 384
DB 301 CTTGGGCAAGCCCTCTGCGCTCTCTCTCGGCACTCTCGGCGCGCCGCGCGAGCTCT 360
QY 385 TGGGGAGAGTCCATCTGTTCCGCGAGAGCCCGCGGCAATACAGATACCTTACGGG 444
DB 361 TGGGGAGAGTCCATCTGTTCCGCGAGAGCCCGCGGCAATACAGATACCTTACGGG 420
QY 445 CAAGTGAGCCAGAGGCGCTTCCCAAGCAGTACCCCTCTTCCGCGCCCTCGCGAGTG 504
DB 421 CAAGTGAGCCAGAGGCGCTTCCCAAGCAGTACCCCTCTTCCGCGCCCTCGCGAGTG 480
QY 505 GTCTTCGCTGCTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTA 564
DB 481 GTCTTCGCTGCTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTA 540
QY 565 CGTCAGTAAGCGGCTGGCGACTTTGGGAGCGCGGGAGGCGCTGGGCGCTGATGAAGGA 624
DB 541 CGTCAGTAAGCGGCTGGCGACTTTGGGAGCGCGGGAGGCGCTGGGCGCTGATGAAGGA 600
QY 625 GATCGAGGCGCGGGAGGCGCTGCAGAGCGTGCAGAGGTGTTTCGCGCCCGCCGCT 684
DB 601 GATCGAGGCGCGGGAGGCGCTGCAGAGCGTGCAGAGGTGTTTCGCGCCCGCCGCT 660
QY 685 CCCAGCGGCGCGGGCAGACGCTCGGCGAGCTGGAGGTGCAGCGAGGCACTCGCTGGT 744
DB 661 CCCAGCGGCGCGGGCAGACGCTCGGCGAGCTGGAGGTGCAGCGAGGCACTCGCTGGT 720
QY 745 CTCGTTTGTGGTGGGATCTGTCGCCAGCCCGCTGCTGGTGGGCTGGACAGCCCTGGA 804
DB 721 CTCGTTTGTGGTGGGATCTGTCGCCAGCCCGCTGCTGGTGGGCTGGACAGCCCTGGA 780
QY 805 CCTGCTGGAGCGGCGGTTGGGGAACAGCGCGGCTGAGCTGTACCCCTACGACCG 864
DB 781 CCTGCTGGAGCGGCGGTTGGGGAACAGCGCGGCTGAGCTGTACCCCTACGACCG 840
QY 865 CGGAGCGGACGCGGCTTACCTTCTCTCTCCGCCAACCTTCCGCCACCATCCCGAGGACAC 924

DB 841 CGGGAGCGGACAGCGGCTTCACTTCTCCTCCCAACTTCGCCACCATCCCGAGGACAC 900
QY 925 GGTGACCGAGATTAACCTCTCTCTCCAGCACCCGCGCAACTCTTCTACTACCGCG 984
DB 901 GGTGACCGAGATTAACCTCTCTCTCCAGCACCCGCGCAACTCTTCTACTACCGCG 960
QY 985 GCTGAAGGCGCTTCCCTCCCATCGCCAGGGTGACACTGGTGGGCTCGACAGAGCCCCAG 1044
DB 961 GCTGAAGGCGCTTCCCTCCCATCGCCAGGGTGACACTGGTGGGCTCGACAGAGCCCCAG 1020
QY 1045 GGCCTTTCATCCCTCCCGCCCCAGTCTCTGCCAGAGGACAATGAGATTGTAGACAGCGC 1104
DB 1021 GGCCTTTCATCCCTCCCGCCCCAGTCTCTGCCAGAGGACAATGAGATTGTAGACAGCGC 1080
QY 1105 CTCAGTTCCAGAAACGCGCTGGACTGGAGGCTCCCTGTGTGTCTCTGGGAGCTGTG 1164
DB 1081 CTCAGTTCCAGAAACGCGCTGGACTGGAGGCTCCCTGTGTGTCTCTGGGAGCTGTG 1140
QY 1165 CGGAGGCCACTGTGGAGGCTCGGGACCAAGAGCAGGACTCGCTACGTCCGGGTCCAGCC 1224
DB 1141 CGGAGGCCACTGTGGAGGCTCGGGACCAAGAGCAGGACTCGCTACGTCCGGGTCCAGCC 1200
QY 1225 CGCCAAACAGGAGGCCCTTCCCGGAGCTCGAAGAGAGGCTGAGTGCCTCCCTGATAA 1284
DB 1201 CGCCAAACAGGAGGCCCTTCCCGGAGCTCGAAGAGAGGCTGAGTGCCTCCCTGATAA 1260
QY 1285 CTCGCTTAAGACCAGAGCCCGCCAGCCCTTGGGGCCCCCGGAGGCAATGGGTGTGCGG 1344
DB 1261 CTCGCTTAAGACCAGAGCCCGCCAGCCCTTGGGGCCCCCGGAGGCAATGGGTGTGCGG 1320
QY 1345 GGCCTCTGTGAGGCTCATGTCTCAGCGCGCCGAGGCGACAGGGGTTTCGCGCTGCTCC 1404
DB 1321 GGCCTCTGTGAGGCTCATGTCTCAGCGCGCCGAGGCGACAGGGGTTTCGCGCTGCTCC 1380
QY 1405 TGACCCGCGGTGAGGCGCGCGCCAGCCATCTCTGCACCTGAAGGGCCCTCTGTTGGCCCGGAC 1464
DB 1381 TGACCCGCGGTGAGGCGCGCGCCAGCCATCTCTGCACCTGAAGGGCCCTCTGTTGGCCCGGAC 1440
QY 1465 GGCATTTGGGAACAGGCTCTCTCTCGCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCT 1584
DB 1441 GGCATTTGGGAACAGGCTCTCTCTCGCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCT 1500
QY 1525 GTCTGCTCTCAGGCTCTCTCTCTCGCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCT 1584
DB 1501 GTCTGCTCTCAGGCTCTCTCTCTCGCTGAGGATAAAGTCAATCCCAAGGCTCCAGCTACTCT 1560
QY 1585 AAATTAATGCTCTCTTATAAGTTATTGCTGCTCCAGGAGATTGCTCTCATGCTCCAGGGG 1644
DB 1561 AAATTAATGCTCTCTTATAAGTTATTGCTGCTCCAGGAGATTGCTCTCATGCTCCAGGGG 1620
QY 1645 CTTGGCTCCACGCTGGTTGCAGATACCTCAGACCTGGTGTCTAGGCTGTGCTGAGGCCCA 1704
DB 1621 CTTGGCTCCACGCTGGTTGCAGATACCTCAGACCTGGTGTCTAGGCTGTGCTGAGGCCCA 1680
QY 1705 CTCCTCCGAGGCGCATCCAAAGCGGGGCGACCTTGAGAAGTGAATAAATGGGCGGTTTC 1764
DB 1681 CTCCTCCGAGGCGCATCCAAAGCGGGGCGACCTTGAGAAGTGAATAAATGGGCGGTTTC 1740
QY 1765 GGAAGGCTCAGTGTGTTTCCATGTTATGATCTCTCTCGGTTTGAATAAAGACTATCTCTGT 1824
DB 1741 GGAAGGCTCAGTGTGTTTCCATGTTATGATCTCTCTCGGTTTGAATAAAGACTATCTCTGT 1800
QY 1825 TGTCTCAAAAAA 1840
DB 1801 TGTCTCAAAAAA 1816

RESULT 8
AAK94244
ID AAK94244 standard; cDNA; 1798 BP.
XX
AC AAK94244;

XX 06-NOV-2001 (first entry)
XX Human full-length cDNA, SEQ ID NO: 2846.
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
OS EP1130094-A2.
XX
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-01183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR P-PSDB; AAM93324.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 2846; 1380pp + sequence listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 1798 BP; 291 A; 619 C; 566 G; 322 T; 0 other;
SQ
Query Match 97.5%; Score 1793.2; DB 22; Length 1798;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1795; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 32 GAAGAGGGTGATCCGACCCGGGAGGTGCGTGGGCGAGGCGAGTTGGAAAGCGGCAGC 91
DB 1 GAAGAGGGTGATCCGACCCGGGAGGTGCGTGGGCGAGGCGAGTTGGAAAGCGGCAGC 60
QY 92 CCCCGCGCCCGCGAGCCCT 151
DB 61 CCCCGCGCCCGCGAGCCCT 120
QY 152 GAGGCGAGCGCGTGAGATCGAAGACAGAGAACTGAGAGCTCATTTGGCGGCGCCCGG 211
DB 121 GAGGCGAGCGCGTGAGATCGAAGACAGAGAACTGAGAGCTCATTTGGCGGCGCCCGG 180
QY 212 GCGCGCGCTCGGGCTTAATAGGAGCTCCGGCTCTGGCTGGAGCCCGACCGCTGCCG 271
DB 181 GCGCGCGCTCGGGCTTAATAGGAGCTCCGGCTCTGGCTGGAGCCCGACCGCTGCCG 240
QY 272 CCGCGCTCCCGTGTCT 331
DB 241 CCGCGCTCCCGTGTCT 300
QY 332 AGGCGCTCTGGCT 391
DB

DB 301 AAGGCCCTCTGCGCT 360
QY 392 GAGTCCATCTGTTCCGCCAGAGCCCGGCAAAATACAGCATCACCTTTCACGGCAAGTGG 451
DB 361 GAATCCATCTGTTCCGCCAGAGCCCGGCAAAATACAGCATCACCTTTCACGGCAAGTGG 420
QY 452 AGCCAGAGCGGCTTCCCAAGCAGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 511
DB 421 AGCCAGAGCGGCTTCCCAAGCAGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 512 CTGCTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTCA 571
DB 481 CTGCTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTCA 540
QY 572 AACGGGCTCGCGACTTTCGGAGCGCGGAGGCTGCGGCTGATGAAGAGATCGAG 631
DB 541 AACGGGCTCGCGACTTTCGGAGCGCGGAGGCTGCGGCTGATGAAGAGATCGAG 600
QY 632 GCGGCGGGGAGGCGCTGCAGAGCGGTGCAGAGGTGTTTTTCGGCGCCCGCTCCCAAG 691
DB 601 GCGGCGGGGAGGCGCTGCAGAGCGGTGCAGAGGTGTTTTTCGGCGCCCGCTCCCAAG 660
QY 692 GGCACCGGCGACAGCTCGCGGAGCTGAGAGTGCAGCGAGGACTCGCTGGTCTCGTT 751
DB 661 GGCACCGGCGACAGCTCGCGGAGCTGAGAGTGCAGCGAGGACTCGCTGGTCTCGTT 720
QY 752 GTGGTGCATCTGTCGCCAGCCCGACTGTTGCTGGGCTGGACAGCTGACAGCTGTGC 811
DB 721 GTGGTGCATCTGTCGCCAGCCCGACTGTTGCTGGGCTGGACAGCTGACAGCTGTGC 780
QY 812 GACGGGACCGTTGGCGGGAACAGCGCGCTGGACCTGTACCCCTACGACGCCGGGAGC 871
DB 781 GACGGGACCGTTGGCGGGAACAGCGCGCTGGACCTGTACCCCTACGACGCCGGGAGC 840
QY 872 GACAGGGGTTCACTTCT 931
DB 841 GACAGGGGTTCACTTCT 900
QY 932 GAGATAACCT 991
DB 901 GAGATAACCT 960
QY 992 GGCCTGCT 1051
DB 961 GGCCTGCT 1020
QY 1052 ATCCCTCCCGCCCGAGTCT 1111
DB 1021 ATCCCTCCCGCCCGAGTCT 1080
QY 1112 CCAGAAACCGCTGGACTGCGAGGTCTCCCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1171
DB 1081 CCAGAAACCGCTGGACTGCGAGGTCTCCCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1172 CACTGTGGAGGCTCGGGACCAAGAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1231
DB 1141 CACTGTGGAGGCTCGGGACCAAGAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
QY 1232 AACGGAGGCGCTCGCCCGAGCTCGAAGAGGCTGAGTGGTCCCTGATTAACCTCGCTC 1291
DB 1201 AACGGAGGCGCTCGCCCGAGCTCGAAGAGGCTGAGTGGTCCCTGATTAACCTCGCTC 1260
QY 1292 TAAGACAGAGCCCGCAGCCCTGGGGCCCGCCGAGGCATGGGGTGTCTGGGGGCTCT 1351
DB 1261 TAAGACAGAGCCCGCAGCCCTGGGGCCCGCCGAGGCATGGGGTGTCTGGGGGCTCT 1320
QY 1352 GTGCGAGGCTCATGCTGCGAGGCGCGAGGCGACAGGGGTTTTTCGCGCTCTCTCTGACC 1411
DB 1321 GTGCGAGGCTCATGCTGCGAGGCGCGAGGCGACAGGGGTTTTTCGCGCTCTCTCTGACC 1380
QY 1412 GGTGAGGCGCGCGGACCATCTCTGCTGAGAGGCGCTCTGTTGGCCCGGACGGGCATT 1471
DB 1381 GGTGAGGCGCGCGGACCATCTCTGCTGAGAGGCGCTCTGTTGGCCCGGACGGGCATT 1440

QY 1472 GGGAAACAGCCCTCCTCTTCCCAACCTTGTCTTTAGGGGCCCGGTGTCCCGTGTCT 1531
|||||
Db 1441 GGGAAACAGCCCTCCTCTTCCCAACCTTGTCTTTAGGGGCCCGGTGTCCCGTGTCT 1500
|||||
QY 1532 CTCAGCCTCCTCCTCTCTGAGGATAAAGTCATCCCAAGGCTCCAGCTACTCTAAATTAT 1591
|||||
Db 1501 CTCAGCCTCCTCCTCTCTGAGGATAAAGTCATCCCAAGGCTCCAGCTACTCTAAATTAT 1560
|||||
QY 1592 GTCTCTTATAAGTTATTCCTCTCCAGGAGATGTCTTCATCTGTCAGGGGCTGGCT 1651
|||||
Db 1561 GTCTCTTATAAGTTATTCCTCTCCAGGAGATGTCTTCATCTGTCAGGGGCTGGCT 1620
|||||
QY 1652 CCCAGCTGGTTCAGATACCTCAGACCTGGTGTCTTAGGCTGTGTGAGCCCACTCTCC 1711
|||||
Db 1621 CCCAGCTGGTTCAGATACCTCAGACCTGGTGTCTTAGGCTGTGTGAGCCCACTCTCC 1680
|||||
QY 1712 GAGGGGCGATCAAGGGGGGCGCACTTGAGAGTGAATAAATGGGGCGGTTCGGAAGCG 1771
|||||
Db 1681 GAGGGGCGATCAAGGGGGGCGCACTTGAGAGTGAATAAATGGGGCGGTTCGGAAGCG 1740
|||||
QY 1772 TCAGTGTTCCTCATGTTATGATCTCTCTCGGTTTGAATAAAGACTATCTCTGTGCTC 1829
|||||
Db 1741 TCAGTGTTCCTCATGTTATGATCTCTCTCGGTTTGAATAAAGACTATCTCTGTGCTC 1798
|||||

RESULT 9
AAF90566
ID AAF90566 standard; DNA; 1785 BP.
XX
AC AAF90566;
XX
DT 22-AUG-2001 (first entry)
XX
DE Human extracellular matrix protein Rgl DNA.
XX
KW Rgl; human; extracellular matrix protein; prostate cancer;
KW metastasis; tumour; benign prostatic hyperplasia; gene therapy;
KW diagnosis; antitumour; ds.
XX
OS Homo sapiens.
XX
FH Key
FT CDS 296..1294
FT /*tag= a
XX
PN WO200144291-A2.
XX
PD 21-JUN-2001.
XX
PF 15-DEC-2000; 2000WO-US33901.
XX
PR 16-DEC-1999; 99US-0172370.
PR 07-DEC-2000; 2000US-9966561.
XX
XX (SCHD) SCHERING AG.
XX
XX Harkins R, Parkes D, Parry G, Schneider DW, Steinbrecher R;
XX
XX WPI; 2001-398128/42.
DR P-PSDB; AAB82472.
XX
XX Novel human extracellular matrix polypeptide, Rgl, useful in research,
PT diagnosis and treatment of metastasis such as prostate cancer
XX
XX Claim 8; Fig 1; 69pp; English.
XX
XX The present sequence is that of a polynucleotide, designated rgl,
CC encoding novel human extracellular matrix polypeptide Rgl (see
CC AAB82472). The full-length coding sequence was deduced from a
CC clone identified from a database screening that was found in the
CC category of cell adhesion molecules and described as a homologue
CC of f-spondin. Rgl is expressed in prostate tissue and may be

over-expressed in prostate tumours. It shows homology to a
superfamily of extracellular matrix proteins encoded by the
mindin/f-spondin genes. The invention provides human Rgl
polypeptides, polynucleotides encoding them (especially encoding
amino acids 1-331, 21-331 or 27-331 of Rgl), methods for
producing the polypeptides, expression vectors and genetically
engineered host cells for expression of the polypeptides. It
also provides methods for utilizing the polynucleotides and
polypeptides in research, diagnosis and therapeutic applications.
Thus, rgl polynucleotides can be used in the production of
recombinant proteins, in methods for assessing Rgl expression
in cells, and in the development of ribozyme and antisense
oligonucleotides useful for administration e.g. to a human
patient for treatment of a disease such as prostate cancer or
benign prostatic hyperplasia, which is alleviated by decreasing
the level of Rgl activity.
XX
SQ Sequence 1785 BP; 300 A; 614 C; 570 G; 301 T; 0 other;
Query Match 95.7%; Score 1761.2; DB 22; Length 1785;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1763; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 9 GGGGTGCGCAGCACTGCGAGGGGAAGAGGGTGATCCGACCCGGGGAAGTCTCGTGGCA 68
|||||
Db 6 GGGGTGCGCAGCACTGCGAGGGGAAGAGGGTGATCCGACCCGGGGAAGTCTCGTGGCA 65
|||||
QY 69 GGGCGAGTTGGAAAGCGGCGCCCGCGCCCGCGAGCCCTTCTCTCTCTCTCTCC 128
|||||
Db 66 GGGCGAGTTGGAAAGCGGCGCCCGCGCCCGCGAGCCCTTCTCTCTCTCTCTCC 125
|||||
QY 129 CAGCTCTATCTGCTCTCTGAGGCCAGGCGCTGCAGCATCGAAGACAGAGGAACT 188
|||||
Db 126 CAGCTCTATCTGCTCTCTGAGGCCAGGCGCTGCAGCATCGAAGACAGAGGAACT 185
|||||
QY 189 GAGGCTCATTTGGCGGCGCCGGGCGCCGCTTCGGGCTTAAATAGAGCTCCGGGCTCT 248
|||||
Db 186 GAGGCTCATTTGGCGGCGCCGGGCGCCGCTTCGGGCTTAAATAGAGCTCCGGGCTCT 245
|||||
QY 249 GGTGGGACCGGCGCTGCGGCGCGGCTGCGGCTGCTCTCTCTCTCTCTCTCTCTCC 308
|||||
Db 246 GGTGGGACCGGCGCTGCGGCGCGGCTGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCC 305
|||||
QY 309 CCAGCCCGCGCGCGCTGCGGCAAGGCGCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCC 368
|||||
Db 306 CCAGCCCGCGCGCGCTGCGGCAAGGCGCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCC 365
|||||
QY 369 CCGCCCGCGCGCGCTCTTTGGGGAGAGTCCATCTGTTCCGCGAGAGCCCGCGGCAATACA 428
|||||
Db 366 CCGCCCGCGCGCGCTCTTTGGGGAGAGTCCATCTGTTCCGCGAGAGCCCGCGGCAATACA 425
|||||
QY 429 GCATCACCTTCAGCGGCAAGTGGAGCCAGAGCGGCTTCCCAAGCAGTACCCCTGTTC 488
|||||
Db 426 GCATCACCTTCAGCGGCAAGTGGAGCCAGAGCGGCTTCCCAAGCAGTACCCCTGTTC 485
|||||
QY 489 GCGCCCTCTGCGAGTGTCTCTGCTCTGCGGCGCGCATAGCTCCGACTACAGCATGT 548
|||||
Db 486 GCGCCCTCTGCGAGTGTCTCTGCTCTGCGGCGCGCATAGCTCCGACTACAGCATGT 545
|||||
QY 549 GGAGGAAGAACCACTACGTACGTACGCGGCTGCGGCACTTTGCGGAGCGCGCGAGCGCT 608
|||||
Db 546 GGAGGAAGAACCACTACGTACGTACGCGGCTGCGGCACTTTGCGGAGCGCGCGAGCGCT 605
|||||
QY 609 GGGCGCTGATGAAGAGATCGAGGCGCGGCGGAGCGCTGCGAGCGTGCAGAGGTGT 668
|||||
Db 606 GGGCGCTGATGAAGAGATCGAGGCGCGGCGGAGCGCTGCGAGCGTGCAGAGGTGT 665
|||||
QY 669 TTTGCGCGCGCGCTCTCTGCTCTGCGGCGCGCATAGCTGCGGCGAGCTGCAGGTGCAGC 728
|||||
Db 666 TTTGCGCGCGCGCTCTCTGCTCTGCGGCGCGCATAGCTGCGGCGAGCTGCAGGTGCAGC 725
|||||
QY 729 GCAGGCACTCGTGTCTCTGTTGTGTGCGCATCTGCTCCCGAGCCCGGCTGTGTCTGG 788
|||||

Db	726	GCAGGCATCGCTGCTCGTTTGTGTGCGCATCGTGCCAGCCCGCATGTGGTTCGTGG	785
Qy	789	GCGTGACAGCCTTGGACCTGTGCGACGGGGACCGTTGGCGGGAACAGCGCGCTGGACC	848
Db	786	GCGTGACAGCCTTGGACCTGTGCGACGGGGACCGTTGGCGGGAACAGCGCGCTGGACC	845
Qy	849	TGTACCCCTTACGACGCCGGGAGCGACAGGGCTTACCTTCTCTCTCCCAACTTTCGCCA	908
Db	846	TGTACCCCTTACGACGCCGGGAGCGACAGGGCTTACCTTCTCTCTCCCAACTTTCGCCA	905
Qy	509	CCATCCGACGAGACGGTGACCGAGATAACGCTCTCTCTCCAGACCCACCGGCCCACT	968
Db	906	CCATCCGACGAGACGGTGACCGAGATAACGCTCTCTCTCCAGACCCACCGGCCCACT	965
Qy	969	CCTTCTACTACCCCGGGCTGAAGGGCCCTGCCCTCCCATCGCCAGGGTGACACTGTGTGGC	1028
Db	966	CCTTCTACTACCCACGGCTGAAGGGCCCTGCCCTCCCATCGCCAGGGTGACACTGTGTGGC	1025
Qy	1029	TGCGACAGAGCCCGAGGGCTTTCATCCTCTCCGGCCCCAGTCTCTGCCAGAGGACAATG	1088
Db	1026	TGCGACAGAGCCCGAGGGCTTTCATCCTCTCCGGCCCCAGTCTCTGCCAGAGGACAATG	1085
Qy	1089	AGATTGTAGACAGCGCTCAGTTCAGAAACCGCGCTGGACTGCGAGGTCTCCCTGTGGT	1148
Db	1086	AGATTGTAGACAGCGCTCAGTTCAGAAACCGCGCTGGACTGCGAGGTCTCCCTGTGGT	1145
Qy	1149	CGTCTGGGACTGTGCGAGGCCACTGTGGAGGCTCGGACCAAGAGGAGACTCGCT	1208
Db	1146	CGTCTGGGACTGTGCGAGGCCACTGTGGAGGCTCGGACCAAGAGGAGACTCGCT	1205
Qy	1209	ACGTCCGGTTCAGCCGCCCAACACGGGAGCCCTGCCCGAGCTCGAAGAAGAGCGTG	1268
Db	1206	ACGTCCGGTTCAGCCGCCCAACACGGGAGCCCTGCCCGAGCTCGAAGAAGAGCGTG	1265
Qy	1269	AGTGCCTCCCTGATAACTCGCTTAAGACACAGAGCCCGCACGCCCTGGGCGCCCGCGA	1328
Db	1266	AGTGCCTCCCTGATAACTCGCTTAAGACACAGAGCCCGCACGCCCTGGGCGCCCGCGA	1325
Qy	1329	GCCATGGGTGTGGGGCTCTGTGTGAGGCTCATGTGTGAGGGCGCGAGGGCACAGGG	1388
Db	1326	GCCATGGGTGTGGGGCTCTGTGTGAGGCTCATGTGTGAGGGCGCGAGGGCACAGGG	1385
Qy	1389	GGTTTCGCGCTGCTCTGACCGGGTGAGGCGCGCCGACCATCTCTGCACTGAAGGCC	1448
Db	1386	GGTTTCGCGCTGCTCTGACCGGGTGAGGCGCGCCGACCATCTCTGCACTGAAGGCC	1445
Qy	1449	CTCTGTGGCGGACGGGCATTTGGAAACAGCCTCTCTTTTCCCAACCTTGCTTCTTA	1508
Db	1446	CTCTGTGGCGGACGGGCATTTGGAAACAGCCTCTCTTTTCCCAACCTTGCTTCTTA	1505
Qy	1509	GGGGCCCCGTGTCCGTCTGCTCTCAGCCTCTCTCTCGACGATAAAGTCAATCCCCA	1568
Db	1506	GGGGCCCCGTGTCCGTCTGCTCTCAGCCTCTCTCTCGACGATAAAGTCAATCCCCA	1565
Qy	1569	AGGCTCCAGCTACTTAAATTTATGCTCTTATAAGTTATTTGCTGCTCCAGGAGTTGTC	1628
Db	1566	AGGCTCCAGCTACTTAAATTTATGCTCTTATAAGTTATTTGCTGCTCCAGGAGTTGTC	1625
Qy	1629	CTTTCATGTCCAGGGCCCTGGCTCCCAAGTGTGTGCAGATACCTTCAGACCTGTGTCTTA	1688
Db	1626	CTTTCATGTCCAGGGCCCTGGCTCCCAAGTGTGTGCAGATACCTTCAGACCTGTGTCTTA	1685
Qy	1689	GGCTGTGCTGAGCCCACTCTCCCGAGGGCGCATCCAAAGCGGGGCCACTTTGAGAAGTAA	1748
Db	1686	GGCTGTGCTGAGCCCACTCTCCCGAGGGCGCATCCAAAGCGGGGCCACTTTGAGAAGTAA	1745
Qy	1749	TAAATGGGCGGTTTCGGAAGCGTCA	1774
Db	1746	TAAATGGGCGGTTTCGGAAGCGTCA	1771

ID	AAF90567	standard; DNA; 1770 BP.
XX		
AC	AAF90567;	
XX		
XX	22-AUG-2001	(first entry)
DT		
XX		
DE	Human extracellular matrix protein RGL DNA.	
XX		
KW	RGL; human; extracellular matrix protein; prostate cancer;	
KW	metastasis; tumour; benign prostatic hyperplasia; gene therapy;	
KW	diagnosis; antitumour; ds.	
XX		
OS	Homo sapiens.	
XX		
XX		
PH	Key	Location/Qualifiers
FT	CDS	296..1294
FT		/*tag= a
XX		
XX	WO200144291-A2.	
PN		
XX	21-JUN-2001.	
PD		
XX		
XX	15-DEC-2000; 2000WO-US33901.	
PF		
XX		
XX	16-DEC-1999; 99US-0172370.	
PR		
PR	07-DEC-2000; 2000US-996561.	
XX		
PA	(SCHD) SCHERING AG.	
XX		
PI	Harkins R, Parkes D, Parry G, Schneider DW, Steinbrecher R;	
XX		
XX	WPI; 2001-398128/42.	
DR	P-PSDB; AAB90567.	
DR		
XX		
XX	Novel human extracellular matrix polypeptide, RGL, useful in research	
PT	diagnosis and treatment of metastasis such as prostate cancer	
PT	disclosure; Fig 4; 69pp; English.	
XX		
XX		
CC	The present sequence is that of a polynucleotide, designated rgl,	
CC	encoding novel human extracellular matrix polypeptide RGL (see	
CC	AA82472). The full-length coding sequence was deduced from a	
CC	clone identified from a database screening that was found in the	
CC	category of cell adhesion molecules and described as a homologue	
CC	of f-spondin. RGL is expressed in prostate tissue and may be	
CC	over-expressed in prostate tumours. It shows homology to a	
CC	superfamily of extracellular matrix proteins encoded by the	
CC	mindin/f-spondin genes. The invention provides human RGL	
CC	polypeptides, polynucleotides encoding them (especially encoding	
CC	amino acids 1-331, 21-331 or 27-331 of RGL), methods for	
CC	producing the polypeptides, expression vectors and genetically	
CC	engineered host cells for expression of the polypeptides. It	
CC	also provides methods for utilizing the polynucleotides and	
CC	polypeptides in research, diagnosis and therapeutic applications.	
CC	Thus, rgl polynucleotides can be used in the production of	
CC	recombinant proteins, in methods for assessing RGL expression	
CC	in cells, and in the development of ribozyme and antisense	
CC	oligonucleotides useful for administration e.g. to a human	
CC	patient for treatment of a disease such as prostate cancer or	
CC	benign prostatic hyperplasia, which is alleviated by decreasing	
CC	the level of RGL activity.	
XX		
SQ	Sequence 1770 BP: 285 A; 614 C; 570 G; 301 T; 0 other;	

Db 66 GGGCGAGTTGGGAAAGCGACGCCCGCCGCGCCCGCCGAGCCCTTCTCCTCTTCTCC 125
QY 129 CAGCTCCTATCTGCTCTCTCGCTGGAGCCAGGCGCTGCGAGCATCGAAGACAGAGGAAC 188
Db 126 CAGCTCCTATCTGCTCTCTCGCTGGAGCCAGGCGCTGCGAGCATCGAAGACAGAGGAAC 185
QY 189 GGAGCCTCATTTGGCGGCGCCGCGGCGGCGCTCGGGCTTAAATAGAGACTCCGGGCTCT 248
Db 186 GGAGCCTCATTTGGCGGCGCCGCGGCGGCGCTCGGGCTTAAATAGAGACTCCGGGCTCT 245
QY 249 GGGTGGGACCCAGCCCTCGCGGCGCGCTCGCGCTCTGCGGGTGATGGAAACCC 308
Db 246 GGGTGGGACCCAGCCCTCGCGGCGCGCTCGCGCTCTGCGGGTGATGGAAACCC 305
QY 309 CCAGCGCGCGCGCGCTTGGCAAGGCGCTCTGCGCTCTCTGCGGCACTCTCGGG 368
Db 306 CCAGCGCGCGCGCGCTTGGCAAGGCGCTCTGCGCTCTCTGCGGCACTCTCGGG 365
QY 369 CGCGCGCGCGCGCTCTGCGGCGAGTCCATCTGTTCCGCCAGAGCGCCCGGCCAATACA 428
Db 366 CGCGCGCGCGCGCTCTGCGGCGAGTCCATCTGTTCCGCCAGAGCGCCCGGCCAATACA 425
QY 429 GCATCACCTTCAGGCGCAAGTGGAGCCAGAGCGCTTCCCAAGCAAGTACCCCGCTTCC 488
Db 426 GCATCACCTTCAGGCGCAAGTGGAGCCAGAGCGCTTCCCAAGCAAGTACCCCGCTTCC 485
QY 489 GCCCGCTCGCAGTGGTCTCTGCTGCTGGGCGCGCATAGCTCCGACTACAGCATGT 548
Db 486 GCCCGCTCGCAGTGGTCTCTGCTGCTGGGCGCGCATAGCTCCGACTACAGCATGT 545
QY 549 GGAGGAAGAACCACTACGTACGTAAAGCGCTCGCGACTTTGCGGAGCGCGCGAGCCCT 608
Db 546 GGAGGAAGAACCACTACGTACGTAAAGCGCTCGCGACTTTGCGGAGCGCGCGAGCCCT 605
QY 609 GGGCGCTGATGAAGAGATCGAGGCGCGGCGGAGCGCTGCGAGCGTGCAGAGGTGT 668
Db 606 GGGCGCTGATGAAGAGATCGAGGCGCGGCGGAGCGCTGCGAGCGTGCAGAGGTGT 665
QY 669 TTTTGGCGCGCGCGCTTCCAGCGGCGGCGAGCGTGGGAGGTTGAGGTCAGC 728
Db 666 TTTTGGCGCGCGCGCTTCCAGCGGCGGCGAGCGTGGGAGGTTGAGGTCAGC 725
QY 729 GCAGCGACTCGTGTGCTCTGTTGTCGCGATCGTCCCGAGCGCGCTGTTGCTGG 788
Db 726 GCAGCGACTCGTGTGCTCTGTTGTCGCGATCGTCCCGAGCGCGCTGTTGCTGG 785
QY 789 GCGTGGAGAGCTGACCTGTGGGCGGCGGAGCGTGGGCGGGAACAGGCGGCGCTGGACC 848
Db 786 GCGTGGAGAGCTGACCTGTGGGCGGCGGAGCGTGGGCGGGAACAGGCGGCGCTGGACC 845
QY 849 TGTACCCCTACGACGCGCGGAGCGGAGCGGCTTACCTTCTCCTCCCGCAACTTCGCCA 908
Db 846 TGTACCCCTACGACGCGCGGAGCGGAGCGGCTTACCTTCTCCTCCCGCAACTTCGCCA 905
QY 909 CCATCCCGCAGGACAGCGGTGACGAGATAACGTCCTCTCCAGCCACCGGCGCAACT 968
Db 906 CCATCCCGCAGGACAGCGGTGACGAGATAACGTCCTCTCCAGCCACCGGCGCAACT 965
QY 969 CCTTCTACTACCGGCGCTGAGGCGCTTCCATCCCTCCCGCGCGCTGCTGCCAGCGGCAATG 1028
Db 966 CCTTCTACTACCGGCGCTGAGGCGCTTCCATCCCTCCCGCGCGCTGCTGCCAGCGGCAATG 1025
QY 1029 TCGCAGAGAGCCCGAGGCGCTTCCATCCCTCCCGCGCGCTGCTGCCAGCGGCAATG 1088
Db 1026 TCGCAGAGAGCCCGAGGCGCTTCCATCCCTCCCGCGCGCTGCTGCCAGCGGCAATG 1085
QY 1089 AGATTGTAGACAGCGCTCAGTTCCAGAAACCGCGCTGCGAGTCCCGCTGCTGTGT 1148
Db 1086 AGATTGTAGACAGCGCTCAGTTCCAGAAACCGCGCTGCGAGTCCCGCTGCTGTGT 1145
QY 1149 CGTCTCGGAGCTGTGCGGAGGCGCACTGTGGAGGCTCGGAGCCAGAGCAGGACTCGT 1208

Db 1146 CGTCTCTGGGACTGTGCGGAGGCCACTGTGGAGGCTCGGAGCAAGAGCAGGACTCGT 1205
QY 1209 AGCTCCGGTCCAGCCCGCAACACAGGAGCCCTGCCCGAGCTCGAAGAGAGGCTG 1268
Db 1206 AGCTCCGGTCCAGCCCGCAACACAGGAGCCCTGCCCGAGCTCGAAGAGAGGCTG 1265
QY 1269 AGTGCCTCCCTGATAACTGCGTCTAAGACAGAGCCCGCAGCCCTTGGGGCCCCCGGA 1328
Db 1266 AGTGCCTCCCTGATAACTGCGTCTAAGACAGAGCCCGCAGCCCTTGGGGCCCCCGGA 1325
QY 1329 GCATGGGTGTGCGGGGCTCTGTGAGGCTCATGCTGAGCGCGCGGAGGCAAGG 1388
Db 1326 GCATGGGTGTGCGGGGCTCTGTGAGGCTCATGCTGAGCGCGCGGAGGCAAGG 1385
QY 1389 GGTTCGCGCTGCTGCTGACCGGCTGAGCGCGGAGGCGCGGAGGCAAGG 1448
Db 1386 GGTTCGCGCTGCTGCTGACCGGCTGAGCGCGGAGGCGCGGAGGCAAGG 1445
QY 1449 CTCTGTGTCGCGCAGCGCATTTGGAAACAGCTCTCTCTTCCCAACCTTCTTCTTA 1508
Db 1446 CTCTGTGTCGCGCAGCGCATTTGGAAACAGCTCTCTCTTCCCAACCTTCTTCTTA 1505
QY 1509 GGGGCCCCGCTGCTGCTGCTGCTGAGGCTCTCTGAGGATAAAGTCAATCCCA 1568
Db 1506 GGGGCCCCGCTGCTGCTGCTGCTGAGGCTCTCTGAGGATAAAGTCAATCCCA 1565
QY 1569 AGGCTCCAGCTACTCTAAATATGCTCTCTTATAAGTATTGCTGCCAGGAGATTG 1628
Db 1566 AGGCTCCAGCTACTCTAAATATGCTCTCTTATAAGTATTGCTGCCAGGAGATTG 1625
QY 1629 CTTTCATGTCGAGGCGCTGGCTCCACGCTGGTTGCGAGATACCTCAGACCTGCTCTA 1688
Db 1626 CTTTCATGTCGAGGCGCTGGCTCCACGCTGGTTGCGAGATACCTCAGACCTGCTCTA 1685
QY 1689 GCGTGTGCTGAGCCACTCTCCGAGGCGGCGATCCAAAGGGGGGCGCTTGAGAGTGAA 1748
Db 1686 GCGTGTGCTGAGCCACTCTCCGAGGCGGCGATCCAAAGGGGGGCGCTTGAGAGTGAA 1745
QY 1749 TAAATGGGCGGTTTCGGAAGCGTC 1773
Db 1746 TAAATGGGCGGTTTCGGAAGCGTC 1770

RESULT 11
AAD20094
ID AAD20094 standard; cDNA; 1779 BP.
XX
AC AAD20094;
XX
DT 03-JAN-2002 (first entry)
XX
DE Human NPG-1 cDNA.
KW Human; NPG-1; cytostatic; gene therapy; tumour; prostrate cancer; lcs;
KW Linker Capture Subtraction; genetic alteration; nerve cell growth; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 236..1228
FT /*tag= a
FT /product= *Human NPG-1 protein*
XX
PN US6287777-B1.
XX
PD 11-SEP-2001.
XX
PF 10-AUG-1999; 99US-0371696.
XX
PR 10-MAY-1996; 96US-0644326.
PR 11-FEB-1998; 98US-0022238.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

QY 1801 CGTTTGAATAAAGACTATCTGTTGGCTCAAAAAA 1840
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 1737 CGTTTGAATAAAGATACTCTGTTGCTCAAAAAA 1776
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

RESULT 12
AAC59794
ID AAC59794 standard; DNA; 1718 BP.
XX AAC59794;
AC AAC59794;
XX
XX 26-JAN-2001 (first entry)
XX
XX Human secreted protein encoding DNA clone vp22 1.

XX Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
KW systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
KW haematopoiesis regulation; tissue regrowth; wound healing; haemophilia;
KW Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
KW contraceptive; infection; growth inhibition; hyperproliferative disorder;
KW psoriasis; ds.
XX
OS Homo sapiens.
XX
XX WO2000055375-A1.
XX
XX 21-SEP-2000.
PD
XX
XX 17-MAR-2000; 2000WO-US07285.
XX
XX 17-MAR-1999; 99US-0124808.
PR
XX 17-MAR-1999; 99US-0124916.
PR
XX 17-AUG-1999; 99US-0149639.
PR
XX 01-OCT-1999; 99US-0157247.
PR
XX 29-NOV-1999; 99US-0167824.
PR
XX 15-FEB-2000; 2000US-0182711.
XX
XX (ALPH-) ALPHAGENE INC.
XX

PI Valenzuela D, Yuan O, Hoffman H, Hall J, Raplejo P;
XX
XX WPI; 2000-638211/61.
XX
XX P-PSDB; AAB34693.
DR

XX Novel proteins and polypeptides useful for the treatment of e.g
PT multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
PT cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and
PT ulcers
XX
XX Claim 22; Page 394-395; 493pp; English.
PS

XX This invention relates to 59 human secreted proteins and the nucleotide
CC sequences encoding them. Sequences AAC59788-C59846 and AAB34687-B34745
CC represent the proteins and their encoding nucleotide sequences, and
CC sequences AAB34746-B34771 represent fragments of the proteins. Probes
CC for the DNA sequences are represented by sequences AAC59847-C59596. The
CC proteins exhibit neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antianemic, nootropic, antiparkinsonian,
CC cerebroprotective, haemostatic, vulnerrary, cytostatic, antipsoriatic,
CC antibacterial, virucide, and fungicide activity. The proteins and
CC nucleotide sequences are useful as nutritional sources or supplements
CC and in research. The proteins are useful for treating immune deficiency
CC and disorders, which may be genetic or resulting from infections,
CC autoimmune disorders such as multiple sclerosis, systemic lupus
CC erythematosus, rheumatoid arthritis, and for treating myeloid or lymphoid
CC cell deficiencies such as anaemias by regulating haematopoiesis. The
CC proteins are also useful in compositions for bone, cartilage, tendon,
CC ligament and/or nerve tissue growth or regeneration, for wound healing,
CC tissue repair and replacement and in the treatment of wounds, incisions
CC and ulcers. Other uses include in the treatment of central and
CC peripheral nervous system and neuropathies such as Alzheimer's and
CC Parkinson's diseases and Shy-Drager syndrome, and mechanical and

CC traumatic disorders, such as spinal cord disorders, head trauma and
CC stroke. The proteins may also be used as a contraceptive, and for
CC treating coagulation disorders such as haemophilias. The protein and
CC nucleotide sequences with cadherin activity are useful for treating
CC cancer. Other uses for the protein include for inhibiting the growth,
CC infection or function of, or killing, infectious agents such as bacteria,
CC virus, fungi and other parasites, for effecting biorhythms or cardiac
CC such as height, weight, hair colour, effecting biorhythms or cardiac
CC cycles or rhythms, effecting metabolism, catabolism, anabolism,
CC processing, utilization, storage or elimination of dietary fat, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors, effecting
CC behavioural characteristics, providing analgesic effects and for treating
CC hyperproliferative disorders such as psoriasis.
XX
XX Query Match 91.6%; Score 1686.2; DB 21; Length 1718;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1699; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 138 TCTGCTCTCGCTGGAGGCGAGCGCTGCAGCATCGAAGAGGAGAACTGGAGCTCA 197
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 2 TCTGCTCTCGCTGGAGGCGAGCGCTGCAGCATCGAAGAGGAGAACTGGAGCTCA 61
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 198 TTGGCGCGCGCGCGCGCGCTCGGGCTTAAATAGAGCTCCCGGCTCTGGCTGGGAC 237
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 62 TTGGCGCGCGCGCGCGCGCTCGGGCTTAAATAGAGCTCCCGGCTCTGGCTGGGAC 121
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 258 CCGACCGCTGCGCGCGCGCTCCCGCTGCTCTGCGGGGTGATGAAAAACCCAGCCCGG 317
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 122 CCGACCGCTGCGCGCGCGCTCCCGCTGCTCTGCGGGGTGATGAAAAACCCAGCCCGG 181
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 318 CCGCGCGCTTGGGCAAGGCGCTCTGGGCTCTCTCTGCGGCTCTGCGGCGCGCGCGG 377
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 182 CCGCGCGCTTGGGCAAGGCGCTCTGGGCTCTCTCTGCGGCTCTGCGGCGCGCGCGG 241
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 378 AGCCTCTTGGGGGAGAGTCCATCTGTTCCGCCAGAGCCCGGCCAAATACAGCATCACCT 437
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 242 AGCCTCTTGGGGGAGAGTCCATCTGTTCCGCCAGAGCCCGGCCAAATACAGCATCACCT 301
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 438 TCACGGGCAAGTGGAGCCAGAGCGGCTTCCCAAGCAGTACCCCTGTTCCGCGCCCGCTG 497
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 302 TCACGGGCAAGTGGAGCCAGAGCGG -CTTCCCAAGCAGTACCCCTGTTCCGCGCCCGCTG 360
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 498 CCGAGTGGTCTTCGCTGCTGGGGCGCGCGCATAGCTCCGACTACAGCATGTGGAGAAAGA 557
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 361 CCGAGTGGTCTTCGCTGCTGGGGCGCGCGCATAGCTCCGACTACAGCATGTGGAGAAAGA 420
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 558 ACCAGTACCTCAGTAACGGGCTGCGGCTTTTCCGAGCGCGCGGAGGCTGGGCGCTGA 617
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 421 ACCAGTACCTCAGTAACGGGCTGCGGCTTTTCCGAGCGCGCGGAGGCTGGGCGCTGA 480
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 618 TGAAGAGATCGAGGCGCGGGGAGGCGCTGCAGAGCGTGCAGAGTGTGTTTCGGGCGC 677
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 481 TGAAGAGATCGAGGCGCGGGGAGGCGCTGCAGAGCGTGCAGAGCGTGTGTTTCGGGCGC 540
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 678 CCGCGCTTCCCGAGCGCCCGGCGAGACCTTCGCGGAGCTGGAGGTGCAGCGAGGACT 737
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 541 CCGCGCTTCCCGAGCGCCCGGCGAGACGCTGCGGAGCTGGAGGTGCAGCGGAGGACT 600
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 738 CGCTGCTCTCGTTTGGTGGCGCATCGTCCAGCGCCGACTGTTCTGTTGGGCGTGGACA 797
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 601 CGCTGCTCTCGTTTGGTGGCGCATCGTCCAGCGCCGACTGTTCTGTTGGGCGTGGACA 660
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 798 GCCTGGACCTGTGCGACGGGGACCGGTTGGCGGGAACAGCGGCGCTGGACCTGTACCCCT 857
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 661 GCCTGGACCTGTGCGACGGGGACCGGTTGGCGGGAACAGCGGCGCTGGACCTGTACCCCT 720
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 858 AGACCGCGGGGAGGAGCGGCTTCACTCTCTCTCCCGCAACTTCGCCACCATCCCGC 917
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 721 AGACCGCGGGGAGGAGCGGCTTCACTCTCTCTCCCGCAACTTCGCCACCATCCCGC 780
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
QY 918 AGSACACGCTGACCGGAGATAACGCTCTCTCCAGCCAGCCCGGCAACTCTCTTACT 977
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```
Db 781 AGACACGGTGAACGAGTAAAGTCTCTCTCCACCCACCGGCAACTCCTTACT 840
QY ACCCGGGGTGAAGGCGCTTCCATCCGACGAGGCTGACACTGCGCTGCGACAGA 1037
Db 841 ACCCGGGGTGAAGGCGCTTCCATCCGACGAGGCTGACACTGCGCTGCGACAGA 900
QY GCGCCAGGCGCTTCACTCCCTCCCGCCCGAGTCTGCGCCAGGAGCAATGAGATTGAG 1097
Db 901 GCGCCAGGCGCTTCACTCCCTCCCGCCCGAGTCTGCGCCAGGAGCAATGAGATTGAG 960
QY ACAGCGCTTCACTCCAGAAACCGCTGGACTGCGAGGCTCTCCCTGTGCTCTCTGGG 1157
Db 961 ACAGCGCTTCACTCCAGAAACCGCTGGACTGCGAGGCTCTCCCTGTGCTCTCTGGG 1020
QY GACTGTGCGGAGGCGCTTCACTCCCTCCCGCCCGAGTCTGCGCCAGGAGCAATGAGATTGAG 1217
Db 1021 GACTGTGCGGAGGCGCTTCACTCCCTCCCGCCCGAGTCTGCGCCAGGAGCAATGAGATTGAG 1080
QY TCCAGCGCCGCAACACGAGGCGCTTCCCTGCGCCAGGAGCAATGAGATTGAGTCC 1277
Db 1081 TCCAGCGCCGCAACACGAGGCGCTTCCCTGCGCCAGGAGCAATGAGATTGAGTCC 1140
QY CTGATTAATGCGTCTTAAGACAGAGCGCCGCGAGCCCTGCGGCGCCCGAGGAGTGGGG 1337
Db 1141 CTGATTAATGCGTCTTAAGACAGAGCGCCGCGAGCCCTGCGGCGCCCGAGGAGTGGGG 1200
QY TGTGCGGCGCTCTCTGTCAGGCTTCACTGTCAGGCGCGGAGGCAACAGGCTTTCGCG 1397
Db 1201 TGTGCGGCGCTCTCTGTCAGGCTTCACTGTCAGGCGCGGAGGCAACAGGCTTTCGCG 1260
QY CTGCTCTGACCGCGTGTAGGCGCGCGGAGGCTTCTGCTGAGGCGCTTTCGCTGG 1457
Db 1261 CTGCTCTGACCGCGTGTAGGCGCGCGGAGGCTTCTGCTGAGGCGCTTTCGCTGG 1320
QY CCGGACGCGCATTGGAAACAGGCTTCTCTTCCCAACCTTCTTAGGCGCGCC 1517
Db 1321 CCGGACGCGCATTGGAAACAGGCTTCTCTTCCCAACCTTCTTAGGCGCGCC 1380
QY GTGTCGCGCTGCTCTGTCAGGCTTCTCTCTGTCAGGATAAAGTCAATCCCAAGGCTCCAG 1577
Db 1381 GTGTCGCGCTGCTCTGTCAGGCTTCTCTCTGTCAGGATAAAGTCAATCCCAAGGCTCCAG 1440
QY CTACTCTAAATATGCTCTCTTATAAGTATGCTGCTCCAGGAGATTCTCTTCACTGT 1637
Db 1441 CTACTCTAAATATGCTCTCTTATAAGTATGCTGCTCCAGGAGATTCTCTTCACTGT 1500
QY CCAGGCGCGCTGCTCTCCAGGCGCATCCCAAGCGGGCGCACTTGAGAGTGAATAAATGGGG 1697
Db 1501 CCAGGCGCGCTGCTCTCCAGGCGCATCCCAAGCGGGCGCACTTGAGAGTGAATAAATGGGG 1560
QY GAGCCGCTCTCCAGGCGCATCCCAAGCGGGCGCACTTGAGAGTGAATAAATGGGG 1757
Db 1561 GAGCCGCTCTCCAGGCGCATCCCAAGCGGGCGCACTTGAGAGTGAATAAATGGGG 1620
QY CGGTTTCGGAAGCGTCAGTGTTCATGATTTATGATCTCTGCGTTTGAATAAAGACTA 1817
Db 1621 CGGTTTCGGAAGCGTCAGTGTTCATGATTTATGATCTCTGCGTTTGAATAAAGACTA 1680
QY TCTCTGTTGCTCAAAAAA 1840
Db 1681 TCTCTGTTGCTCAAAAAA 1703
```

RESULT 13
AAH34981
ID AAH34981 standard; cDNA; 1848 BP.
XX
AC AAH34981;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:2063.

```
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; ss.  
XX  
OS Homo sapiens.  
PN WO200122920-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-US26524.  
XX  
PR 29-SEP-1999; 99US-0157137.  
PR 03-NOV-1999; 99US-0163280.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
DR WPI; 2001-235357/24.  
DR P-PSDB; AAG75576.  
XX  
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
XX Claim 1; Page 3549; 9803pp; English.  
XX  
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB7789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
SQ Sequence 1848 BP; 324 A; 601 C; 578 G; 340 T; 5 other;  
Query Match 91.3%; Score 1680.8; DB 22; Length 1848;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1679; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 158 AGCGCGTCGAGCATCGAAGACAGAGGAGAACTGAGGCTCATTTGGCGGCGCGCGCG 217  
Db 133 AGCGCGTCGAGCATCGAAGACAGAGGAGAACTGAGGCTCATTTGGCGGCGCGCGCG 192  
QY 218 GCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACGCTGCGCGCGCG 277  
Db 193 GCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACGCTGCGCGCGCG 252  
QY 278 TCCCGCTGCTCTGCGGGGTGATGGAANAACCCAGCGCGCGCGCGCGCTTGGGCAAGGCC 337  
Db 253 TCCCGCTGCTCTGCGGGGTGATGGAANAACCCAGCGCGCGCGCGCTTGGGCAAGGCC 312  
QY 338 CTCTGGGCTCTCTCTGCGGCACTCTCGCGCGCGCGCGCGCGCTTGGGCGAGAGTCC 397  
Db 313 CTCTGGGCTCTCTCTGCGGCACTCTCGCGCGCGCGCGCGCGCTTGGGCGAGAGTCC 372  
QY 398 ATCTGTTCGCGCAGAGCGCGCGCGCAATACAGCATCACCTTACGCGGCAAGTGGAGCCAG 457  
Db 373 ATCTGTTCGCGCAGAGCGCGCGCGCAATACAGCATCACCTTACGCGGCAAGTGGAGCCAG 432
```

QY 458 ACGGCTTCCCAAGCAGTACCCCTGTTCGCCGCCCTGCGCAGTGTCTTCGCTGCTG 517
|||||
Db 433 ACGGCTTCCCAAGCAGTACCCCTGTTCGCCGCCCTGCGCAGTGTCTTCGCTGCTG 492
|||||
QY 518 GGGGGCGGCATAGCTCCGACTACAGCATGTGGAGAGAACCCAGTACGTCAAGCGG 577
|||||
Db 493 GGGGGCGGCATAGCTCCGACTACAGCATGTGGAGAGAACCCAGTACGTCAAGCGG 552
|||||
QY 578 CTGCGCGACTTTGCGGAGCGCGGAGGCTGCGGCGCTGATGAAGGAGATCCAGCGGGG 637
|||||
Db 553 CTGCGCGACTTTGCGGAGCGCGGAGGCTGCGGCGCTGATGAAGGAGATCCAGCGGGG 612
|||||
QY 638 GGGGAGCGCTCAGAGCGTGCACAGAGTGTTCGCGCGCCGCGCTCCCGACGCGCAC 697
|||||
Db 613 GGGGAGCGCTCAGAGCGTGCACAGTGTTCGCGCGCCGCGCTCCCGACGCGCAC 672
|||||
QY 698 GGGGAGCGTGGGAGAGTGGAGGTGCAGCGCAGCACTCGCTGGTCTCTGTTGTGGTG 757
|||||
Db 673 GGGGAGCGTGGGAGAGTGGAGGTGCAGCGCAGCACTCGCTGGTCTCTGTTGTGGTG 732
|||||
QY 758 CGCATCGTCCCGACGCCCGACTGGTTCGFGGCGTGGACAGCCTGGACCTGTGCGACGG 817
|||||
Db 733 CGCATCGTCCCGACGCCCGACTGGTTCGFGGCGTGGACAGCCTGGACCTGTGCGACGG 792
|||||
QY 818 GACCGTTGGCGGGAACAGCGCGGCTGGACCTGTACCCCTAGCAGCCCGGACGCGACAGC 877
|||||
Db 793 GACCGTTGGCGGGAACAGCGCGGCTGGACCTGTACCCCTAGCAGCCCGGACGCGACAGC 852
|||||
QY 878 GCGTTCACCTTCTCTCCCGCAACTTTCGCCACCATCCCGCAGGACACGGTGACCGAGATA 937
|||||
Db 853 GCGTTCACCTTCTCTCCCGCAACTTTCGCCACCATCCCGCAGGACACGGTGACCGAGATA 912
|||||
QY 938 AGCTCTCTCTCTCCAGCCACCGCGCAACTCTCTTACTACCGCGGCTGAAGCCCTG 997
|||||
Db 913 AGCTCTCTCTCTCCAGCCACCGCGCAACTCTCTTACTACCGCGGCTGAAGCCCTG 972
|||||
QY 998 CTCTCCATCGCGAGGTGACACTGTGTGCGGTGCGACAGAGCCCGAGGCGCTTCATCCCT 1057
|||||
Db 973 CTCTCCATCGCGAGGTGACACTGTGTGCGGTGCGACAGAGCCCGAGGCGCTTCATCCCT 1032
|||||
QY 1058 CCGGCGCCAGTCTGCGCCAGCAGGGAATGAGATTGTAGACAGCGCCTCAGTTCCAGAA 1117
|||||
Db 1033 CCGGCGCCAGTCTGCGCCAGCAGGGAATGAGATTGTAGACAGCGCCTCAGTTCCAGAA 1092
|||||
QY 1118 AGCGCGCTGAGTGGAGTCTCCCTGTGTGCTGTGCGGACTGTCGAGGCGCACTCT 1177
|||||
Db 1093 AGCGCGCTGAGTGGAGTCTCCCTGTGTGCTGTGCGGACTGTCGAGGCGCACTCT 1152
|||||
QY 1178 GGGAGGCTCGGACCAAGACAGGACTCGCTACGTCCGGTCCAGCCGCCCAACACGGG 1237
|||||
Db 1153 GGGAGGCTCGGACCAAGACAGGACTCGCTACGTCCGGTCCAGCCGCCCAACACGGG 1212
|||||
QY 1238 AGCCCTGCCCCAGTCTGAAGAGAGGCTGAGTGGTCCCTGATTAACCTGCGTCTAAGAC 1297
|||||
Db 1213 AGCCCTGCCCCAGTCTGAAGAGAGGCTGAGTGGTCCCTGATTAACCTGCGTCTAAGAC 1272
|||||
QY 1298 CAGAGCCCGCAGCCCTGCGGCCCGCCAGCCATGGGGTGTGCGGGCTCTGTGCA 1357
|||||
Db 1273 CAGAGCCCGCAGCCCTGCGGCCCGCCAGCCATGGGGTGTGCGGGCTCTGTGCA 1332
|||||
QY 1358 GCTCATGTGCGAGGCGCGAGGACAGGGGTTTCGCGCTGCTCCTGACCGCGGTGAG 1417
|||||
Db 1333 GCTCATGTGCGAGGCGCGAGGACAGGGGTTTCGCGCTGCTCCTGACCGCGGTGAG 1392
|||||
QY 1418 GCGGCGCGACCATCTCTGCATGAAGGCGCCTCTGTGTGGCGCGCACGGGCATTGGGAAA 1477
|||||
Db 1393 GCGGCGCGACCATCTCTGCATGAAGGCGCCTCTGTGTGGCGCGCACGGGCATTGGGAAA 1452
|||||
QY 1478 CAGCGCTCTCTCTCCACACCTTGTCTTAGGGGCGCCCGTGTCCCGTCTGCTCAGC 1537
|||||
Db 1453 CAGCGCTCTCTCTCTCCACACCTTGTCTTAGGGGCGCCCGTGTCCCGTCTGCTCAGC 1512
|||||
QY 1538 CTCCTCTCTCTGAGGATAAAGTTCATCCCAAGGCTCCAGGCTACTCTAAATTTATGCTCC 1597
|||||

Db 1513 CTCCTCTCTCTGAGGATAAAGTTCATCCCAAGGCTCCAGCTACTCTAAATTTATGCTCC 1572
|||||
QY 1598 TTATAAGTTATTGCTGCTCCAGGAGATTGCTCTTCATGTCAGGGGCTGGTCCACAG 1657
|||||
Db 1573 TTATAAGTTATTGCTGCTCCAGGAGATTGCTCTTCATGTCAGGGGCTGGTCCACAG 1632
|||||
QY 1658 TCGTTGCAGATACCTCAGACCTGGTCTCTAGGCTGTGCTGAGCCACCTCTCCGAGGGC 1717
|||||
Db 1633 TCGTTGCAGATACCTCAGACCTGGTCTCTAGGCTGTGCTGAGCCACCTCTCCGAGGGC 1692
|||||
QY 1718 GCATCCAAAGCGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTCGGAAGCGTCAGTG 1777
|||||
Db 1693 GCATCCAAAGCGGGGCCACTTGAGAAGTGAATAAATGGGCGGTTCGGAAGCGTCAGTG 1752
|||||
QY 1778 TTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACATATCTCTGCTCAAAAAA 1837
|||||
Db 1753 TTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACATATCTCTGCTCAAAAAA 1812
|||||
QY 1838 AAA 1840
|||||
Db 1813 AAA 1815
|||||
RESULT 14
AAK94182
ID AAK94182 standard; cDNA; 1669 BP.
XX
AC AAK94182;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2726.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR P-PSDB; AAM93266.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 2726; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods.
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

[illegible]

```
XX 11-JUN-1997; 97US-0049288.
PR 10-APR-1997; 97US-0043421.
XX (ZYMO ) ZYMOGENETICS INC.
PA Sheppard PO;
XX PI
XX WPI: 1998-557522/47.
DR P-PSDB; AAW70589.
XX
PT New zsig25 protein and related nucleic acid, fusion proteins,
PT vectors, transformed cells - and antibodies, involved in modulation
PT of adhesion, used for diagnosis and treatment of prostatic and
PT B-cell tumours, stimulation of haematopoietic cells, treatment of
PT immune deficiency etc.
XX
PS Claim 40; Pages 108-110; 161pp; English.
XX
CC The present sequence encodes a protein designated zsig25. The zsig25
CC protein is an adhesion-modulating protein expressed at very high level
CC in prostate cells. The protein is useful as a diagnostic marker for
CC prostatic carcinoma and B-cell cancers, possibly also for infertility,
CC and as a reagent for separating cancerous and non-cancerous cells.
CC The products may also be used to diagnose or treat Wolf-Hirschhorn
CC syndrome, associated with a deletion in the region of chromosome 4
CC (p16.3) where the zsig25 gene is located.
XX
SQ Sequence 1607 BP; 284 A; 540 C; 494 G; 289 T; 0 other;
Query Match 85.0%; Score 1563.4; DB 19; Length 1607;
Best Local Similarity 99.3%; Pred. No. 3.7e-300;
Matches 1570; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 260 GACCGCTGCGCGCGCTCCCGTCTGCTGCTGCGGGTGATGAAACCCCGCGGC 319
DB 1 GAAATTCGGCTCGAGGGCGCTCCCGTCTGCTGCGGGTGATGAAACCCCGCGGC 60
QY 320 GCGCGCTGGGCAAGCCCTCGCTGCTCTCTGCGCACTCGCGCGCGCGGCCAG 379
DB 61 GCGCGCTGGGCAAGCCCTCGCTGCTCTCTGCGCACTCGCGCGCGCGGCCAG 120
QY 380 CCTCTTGGGGAGAGTCCATCTCTTCCGCGAGAGCCCGGCCAAATACAGCATCCTTC 439
DB 121 CCTCTTGGGGAGAGTCCATCTCTTCCGCGAGAGCCCGGCCAAATACAGCATCCTTC 180
QY 440 ACGGCAAGTGGAGCAGAGCGCTTCCCAAGCAGTACCCCTGTTCGCCGCCCTGCG 499
DB 181 ACGGCAAGTGGAGCAGAGCGCTTCCCAAGCAGTACCCCTGTTCGCCGCCCTGCG 240
QY 500 CAGTGTCTTCGCTGCGTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAC 559
DB 241 CAGTGTCTTCGCTGCGTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAC 300
QY 560 CAGTGTCTCAGTAACGGGCTGCGCGACTTTCGCGAGCGCGCGAGGCGCTGGGCGCTGATG 619
DB 301 CAGTGTCTCAGTAACGGGCTGCGCGACTTTCGCGAGCGCGCGAGGCGCTGGGCGCTGATG 360
QY 620 AAGGAGATCAGGCGCGGGGAGGCGCTGCAGAGCGTGCACAGAGTGTTTCGCGGCC 679
DB 361 AAGGAGATCAGGCGCGGGGAGGCGCTGCAGAGCGTGCACAGAGTGTTTCGCGGCC 420
QY 680 GCGCTCCCCAGCGGCGAGAGCTGCGCGAGCTGGAGGTGCAGCGCGCGACTCG 739
DB 421 GCGCTCCCCAGCGGCGAGAGCTGCGCGAGCTGGAGGTGCAGCGCGCGACTCG 480
QY 740 CTGCTCTCGTTTGTGTGCGCATCTGCGCCAGCCCGACTGGTTCTGTGGCGTGGACAGC 799
DB 481 CTGCTCTCGTTTGTGTGCGCATCTGCGCCAGCCCGACTGGTTCTGTGGCGTGGACAGC 540
QY 800 CTGACCTGTGCGACGGGACCGTTGGCGGAACAGCGCGCGCTGGACCTGTACCCCTAC 859
DB 541 CTGACCTGTGCGACGGGACCGTTGGCGGAACAGCGCGCGCTGGACCTGTACCCCTAC 600
```

Search completed: October 27, 2002, 08:12:45
Job time : 307 secs

```
QY 860 GACGCGGGACGACAGCGGCTTCACCTTCTCTCTCCCAACTTCGCCACCATCCCGCAG 919
DB 601 GACGCGGGACGACAGCGGCTTCACCTTCTCTCTCCCAACTTCGCCACCATCCCGCAG 660
QY 920 GACACGGTGACCGAGATAAGCTCTCTCTCCAGGACACCCGGCCAACTCTTCTACTAC 979
DB 661 GACACGGTGACCGAGATAAGCTCTCTCTCCAGGACACCCGGCCAACTCTTCTACTAC 720
QY 980 CCGCGGCTGAAGCCCTGCTCCATCGCCAGGGTGACACTGGTGGGGTGGCACAGAGC 1039
DB 721 CCGCGGCTGAAGCCCTGCTCCATCGCCAGGGTGACACTGGTGGGGTGGCACAGAGC 780
QY 1040 CCCAGGCGCTTCATCCTCCCGCCAGTCTCTCCAGCAGGACAGCATGATGTTAGAC 1099
DB 781 CCCAGGCGCTTCATCCTCTCCCGCCAGTCTCTCCAGCAGGACAGCATGATGTTAGAC 840
QY 1100 AGCGCCTCAGTTCCAGAAACGCGCTGGACTGGAGGTCTCTCTGTGGTGGTCTCTGGGA 1159
DB 841 AGCGCCTCAGTTCCAGAAACGCGCTGGACTGGAGGTCTCTCTGTGGTGGTCTCTGGGA 900
QY 1160 CTGTGGGAGGCGACTGTGGAGGCTCGGACCAAGACAGAGACTCGCTACGTCCGGGTC 1219
DB 901 CTGTGGGAGGCGACTGTGGAGGCTCGGACCAAGACAGAGACTCGCTACGTCCGGGTC 960
QY 1220 CAGCGCGCCCAACACAGGAGCGCTCCCGCGAGCTCGAAGAAAGAGCTGAGTGGTCCCT 1279
DB 961 CAGCGCGCCCAACACAGGAGCGCTCCCGCGAGCTCGAAGAAAGAGCTGAGTGGTCCCT 1020
QY 1280 GATAACTGCGCTTAAGACACAGAGCGCGCGAGCTGGGGCCCCCGAGGCCATGGGTG 1339
DB 1021 GATAACTGCGCTTAAGACACAGAGCGCGCGAGCTGGGGCCCCCGAGGCCATGGGTG 1080
QY 1340 TCGGGGCGCTCTGTGAGGCTCATGTGTGAGCGGCGAGGGCACAGGGGGTTCGCGCT 1399
DB 1081 TCGGGGCGCTCTGTGAGGCTCATGTGTGAGCGGCGAGGGCACAGGGGGTTCGCGCT 1140
QY 1400 GCTCTCAGCGCGGTGAGCGCGCGAGCATCTCTGCACTGAAGGCCCTCTGGTGGCC 1459
DB 1141 GCTCTCAGCGCGGTGAGCGCGCGAGCATCTCTGCACTGAAGGCCCTCTGGTGGCC 1200
QY 1460 GGCACGGGCAATTGGGAAACAGCGCTCTCTTCCCACTTGTCTTAGGGGCCCCCGT 1519
DB 1201 GGCACGGGCAATTGGGAAACAGCGCTCTCTTCCCACTTGTCTTAGGGGCCCCCGT 1260
QY 1520 GTCCCGTCTCTCAGCGCTCTCTCTCGACAGATAAAGTCAATCCCAAGGCTCCAGCT 1579
DB 1261 GTCCCGTCTCTCAGCGCTCTCTCTCGACAGATAAAGTCAATCCCAAGGCTCCAGCT 1320
QY 1580 ACTCTAAATTTATCTCTCTTATAAGTTATTGCTCTCCAGAGATTGCTCTCATCGTCC 1639
DB 1321 ACTCTAAATTTATCTCTCTTATAAGTTATTGCTCTCCAGAGATTGCTCTCATCGTCC 1380
QY 1640 AGGGGCGTGGCTCCCGAGCTGGTTCGAGATACCTCAGACCTGGTCTTAGGCTGTGCTGA 1699
DB 1381 AGGGGCGTGGCTCCCGAGCTGGTTCGAGATACCTCAGACCTGGTCTTAGGCTGTGCTGA 1440
QY 1700 GCCACTCTCCCGAGGGCGCATCCAGCGGGGCGCACTTGAGAGTGAATAAATGGGGCG 1759
DB 1441 GCCACTCTCCCGAGGGCGCATCCAGCGGGGCGCACTTGAGAGTGAATAAATGGGGCG 1500
QY 1760 GTTTCGGAAGCGTCAGTGGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAGACTATC 1819
DB 1501 GTTTCGGAAGCGTCAGTGGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAGACTATC 1560
QY 1820 TCTGTTGCTCAAAAAA 1840
DB 1561 TCTGTTGATCAAAAAA 1581
```


Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	1692.4	92.0	1779	4	US-09-371-696-1	Sequence 1, Appli
2	1100.2	59.8	1105	2	US-08-799-173A-1	Sequence 1, Appli
3	343.8	18.7	400	1	US-08-644-326-1	Sequence 1, Appli
4	339	18.4	400	4	US-09-022-238-1	Sequence 1, Appli
5	304.8	16.6	506	2	US-08-799-173A-15	Sequence 15, Appli
6	204.2	11.1	316	2	US-08-799-173A-16	Sequence 16, Appli
7	204.2	11.1	316	2	US-08-799-173A-17	Sequence 17, Appli
8	71.6	3.9	4029	1	US-07-862-021B-9	Sequence 9, Appli
9	71.6	3.9	4029	1	US-08-313-288B-9	Sequence 9, Appli
10	71.6	3.9	4029	5	PCF-U933-03164-9	Sequence 9, Appli
11	62.2	3.4	3226	1	US-07-862-021B-11	Sequence 11, Appli
12	62.2	3.4	3226	1	US-08-313-288B-11	Sequence 11, Appli
13	62.2	3.4	3226	5	PCF-U933-03164-11	Sequence 11, Appli
14	57.6	3.1	1155	4	US-08-818-112-12	Sequence 12, Appli
15	57.6	3.1	1155	4	US-08-818-111-12	Sequence 12, Appli
16	57.6	3.1	1155	4	US-09-056-556-12	Sequence 12, Appli
17	57.2	3.1	1816	1	US-07-862-021B-13	Sequence 13, Appli
18	57.2	3.1	1816	1	US-08-313-288B-13	Sequence 13, Appli
19	57.2	3.1	1816	5	PCF-U933-03164-13	Sequence 13, Appli
20	55.4	3.0	1105	2	US-08-799-173A-1	Sequence 1, Appli
21	55	3.0	30001	1	US-08-125-468-1	Sequence 1, Appli
22	55	3.0	30001	2	US-08-474-933-1	Sequence 1, Appli
23	54.4	3.0	933	4	US-09-105-390-43	Sequence 43, Appli
24	54.4	3.0	1008	4	US-09-105-390-59	Sequence 59, Appli
25	54.4	3.0	2810	4	US-09-105-390-6	Sequence 6, Appli
26	52.4	2.8	1779	4	US-09-371-696-1	Sequence 1, Appli
27	51.8	2.8	8625	4	US-08-980-832-1	Sequence 1, Appli

LOCATION: 19..1011

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 19..963

US-08-799-173A-1

Query Match

Best Local Similarity 99.7%; Pred. No. 1e-219;

Matches 1102: Conservative 0: Mismatches

281	Qy	CGTGTCTCTCGCGGTGATGGAACCCAGCCGGCCGCGCCTTGGGCAAGGCCCTC	340
1	Db		
		CGTGTCTCTCGCGGTGATGGAACCCAGCCGGCCGCGCCTTGGGCAAGGCCCTC	
341	Qy	TGGGCTCTCTCTGTGCCACTCTCGGCGCGCGGGCCAGCCTCTTGGGGGAGAGTCCATC	400
61	Db	TGGGCTCTCTCTGTGCCACTCTCGGCGCGCGGGCCAGCCTCTTGGGGGAGAGTCCATC	120
401	Qy	TGTTCCGCGAGAGCCCCGGCCAAATACAGCATCATCCTTACGSGGCAAGTGGAGCCAGAG	460
121	Db	TGTTCCGCGAGAGCCCTGGCCAAATACAGCATCATCCTTACGSGGCAAGTGGAGCCAGAG	180
461	Qy	GCCTTCCCAAGAGTATCCCTCTGTTCCGCCGCCCTCGCAGTGTGCTTTCGCTGCTGGGG	520
181	Db	GCCTTCCCAAGAGTATCCCTCTGTTCCGCCGCCCTCGCAGTGTGCTTTCGCTGCTGGGG	240
521	Qy	GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTACGTACAGTACGCGG	580
241	Db	GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTACGTACAGTACGCGG	300
581	Qy	CGGACTTTTCGGAGGCGCGGAGGCGCTGGGGCTGATGAAGGAGATCAGAGCGCGGGG	640
301	Db	CGGACTTTTCGGAGGCGCGGAGGCGCTGGGGCTGATGAAGGAGATCAGAGCGCGGGG	360
641	Qy	GAGGCGCTGAGAGCTGTACAGAGTGTGTTTCGGCGCCCGCGCTGCCAGCGGACACGGG	700
361	Db	GAGGCGCTGAGAGCTGTACAGCGTGTGTTTCGGCGCCCGCGCTGCCAGCGGACACGGG	420
701	Qy	CAGAGCTCGGCGAGCTGGAGGTGCAGCGACGACACTCGCTGCTCNCGTTTGTGTGGCG	760
421	Db	CAGAGCTCGGCGAGCTGGAGGTGCAGCGACGACACTCGCTGCTCNCGTTTGTGTGGCG	480
761	Qy	ATCGTGCCAGCCCGGACTGGTTTCGTGGGCGTGGAGCGCTGGACCTGTGCAGCGGGGAC	820
481	Db	ATCGTGCCAGCCCGGACTGGTTTCGTGGGCGTGGAGCGCTGGACCTGTGCAGCGGGAC	540
821	Qy	CGTTGGGGGAACAGCGGGCGCTGGAGCTGTATCCCTACAGCCCGGAGCAGCAGCGCG	880
541	Db	CGTTGGGGGAACAGCGGGCGCTGGAGCTGTATCCCTACAGCCCGGAGCAGCAGCGCG	600
881	Qy	TTCACTCTCTCTCCCCCAACTTCGCCACATCCCGCAGGACAGGTGACCCAGAGATAAG	940
601	Db	TTCACTCTCTCTCCCCCAACTTCGCCACATCCCGCAGGACAGGTGACCCAGAGATAAG	660
941	Qy	TCCTCTCTCTCCAGCCCGGCCAACTCCTTCTACTACCCGGGCTGAAGGCCCTGCCT	1000
661	Db	TCCTCTCTCTCCAGCCCGGCCAACTCCTTCTACTACCCGGGCTGAAGGCCCTGCCT	720
1001	Qy	CCCATCGCCAGGTGACACTGTGTGGCTGTGCGGCTGCGACAGAGCCCGAGGCCCTTCAT	1060
721	Db	CCCATCGCCAGGTGACACTGTGTGGCTGTGCGGCTGCGACAGAGCCCGAGGCCCTTCAT	780
1061	Qy	GCCCCAGTCTCTGCCAGCAGGAGCAATGAGATTGTAGACAGCGCCTCAGTCCAGAAAG	1120
781	Db	GCCCCAGTCTCTGCCAGCAGGAGCAATGAGATTGTAGACAGCGCCTCAGTTCAGAAAG	840
1121	Qy	CCGCTGGACTCGAGGTCTCCCTGTGGTGTGCTGTGGGACTGTGCGGAGGCCACTGTGGG	1180
841	Db	CCGCTGGACTCGAGGTCTCCCTGTGGTGTGCTGTGGGACTGTGCGGAGGCCACTGTGGG	900
1181	Qy	AGGCTCGGGACCAAGAGCAGGACTCGCTACGTCCGGGTCCAGCCCCGCCAACACAGGGAG	1240
901	Db	AGGCTCGGGACCAAGAGCAGGACTCGCTACGTCCGGGTCCAGCCCCGCCAACACAGGGAG	960

QY 1090 GATTGTAGACAGCGCTCAGTTCCAGAAAGCGCGCTGGAGTGGAGGTCTCCCTGTGGTC 1149
|||||
Db 236 GATTGTAGACAGCGCTCAGTTCCAGAAAGCGCGCTGGAGTGGAGGTCTCCCTGTGGTC 295
QY 1150 GTCTCTGGGACTGTGGAGGCGCACTGTGGAGGCTCGGAGCAAGAGGAGTCCGCTA 1209
|||||
Db 296 GTCTCTGGGACTGTGGAGGCGCACTGTGGAGGCTCGGAGTCCGCTA 355
QY 1210 CGTCCGGGTCCAGCGCGCAACAGCGGAGCGCCCTGCCCGGAGCT 1254
|||||
Db 356 CGCCCGGTCCAGCGCGCAACAGCGGAGCGCCCTGCCCGGAGCT 400

RESULT 4

US-09-022-238-1
; Sequence 1, Application US/09022238
; Patent No. 6177244
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J. and Yang, Meiheng
; TITLE OF INVENTION: A novel NPG-1 Gene that is differentially expressed in prosta
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/022,238

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/644,326
; FILING DATE: 10-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Maravic-Magovcovic, Ivana
; REGISTRATION NUMBER: P-43,338
; REFERENCE/DOCKET NUMBER: NER-262CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..398

US-09-022-238-1

Query Match 18.4%; Score 339; DB 4; Length 400;
Best Local Similarity 96.3%; Pred. No. 5.3e-62;
Matches 390; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

QY 850 GTACCCCTAGAGCGCGGAGGACAGCGGTTACCTTCTCTCCCAACTTCGCCAC 909
|||||
Db 1 GTACCCCTAGAGCGCGGAGGAC-GACAGCGGTTACCTTCTCTCCCAACTTCGCCAC 59
QY 910 CATCCCGGAGGACAGGTACCGAGATACGTCCTCTCCAGCCACCGCGCAACTC 969
|||||
Db 60 CATCCCGCA-GACACGGTGACCGAGATACGTCCTCTCCAGCCACCGCGCAACTC 118
QY 970 CTCTACTACCGCGGCTGAAGGCGCTGCTCCCAATCGCCAGGTTGACACTGTGTGGGCT 1029

Db 119 CTCTACTACCGCGGCTGAAGGCC--TGCTCCATCGCAGGGTGACACTGTGTGGCT 176
QY 1030 GGCACAGAGCCCGAGGCGCTTCATCCCTCCGCCCCAGTCTCCCGCAGGAGCAATGA 1089
Db 177 GGCACAGAGCCCGAGGCGCTTCATCCCTCCGCCCCAGTCTCCCGCAGCA-GGACAATGC 235
QY 1090 GATTGTAGACAGCGCTCAGTTCCAGAAAGCGCGCTGGAGTGGAGGTCTCCCTGTGGTC 1149
Db 236 GATTGTAGACAGCGCTCAGTTCCAGAAAGCGCGCTGGAGTGGAGGTCTCCCTGTGGTC 295
QY 1150 GTCCTGGGACTGTGGAGGCGCACTGTGGAGGCTCGGAGCAAGAGGAGTCCGCTA 1209
Db 296 GTCCTGGGACTGTGGAGGCGCACTGTGGAGGCTCGGAGTCCGCTA 355
QY 1210 CGTCCGGGTCCAGCGCGCAACAGCGGAGCGCCCTGCCCGGAGCT 1254
Db 356 CGCCCGGTCCAGCGCGCAACAGCGGAGCGCCCTGCCCGGAGCT 400

RESULT 5

US-08-799-173A-15
; Sequence 15, Application US/08799173A
; Patent No. 5871969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG.
; APPLICANT: PATRICK J. DILLON
; TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 11-FEB-1997
; APPLICATION NUMBER: US/08/799,173A
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-799-173A-15

Query Match 16.6%; Score 304.8; DB 2; Length 506;
Best Local Similarity 85.5%; Pred. No. 6.9e-55;
Matches 371; Conservative 0; Mismatches 59; Indels 4; Gaps 4;

QY 293 CGGGTGATGGAACCCAGCCCGCGCGCTGGGAGGCCCTCTGCGCTCTCCTC 352
|||||
Db 6 CGGCAAGAGNNAAACCCAGCCCGCGCTGGGAGGCCCTCTGCGCTCTCCTC 65
QY 353 CTGGCCACTCTCGGGCGCGCGCGAGCTCTTTGGGGAGAGTCCATCTGTTCCGCCAGA 412
|||||
Db 66 CTGGCCACTCTCGGGCGCGCGC-ACCAGCCTCTTTGGGGAGAGTCCATCTNTTCCGCCAGA 124

QY 413 GCCCGGCGAAATACAGCATACCTTACGGGCAAGTGGAGCCAGCGGCTTCCCAAG 472
Db 125 GCCCGGCGAAATACAGCATACCTTACGGGCAAGTGGAGCCAGCGGCTTCCCAAG 184
QY 473 CAGTACCCCTGTTCGCGCCCTCGCGAGTGTCTTCGCTGTGGGCGCGCATAGC 532
Db 185 CAGTACCCCTGTTCGCGCCCTCGCGCA-TGGTNTTCGCTGTGGGCGCGCATAGC 243
QY 533 TCCGACTACAGCATGTGGAGGAAGAACAGTACGTCACTAAGCGGCTGCGGACTTTGG 592
Db 244 TCCGACTACAGCATGTGGAGGAAGAACAGTACGTCACTAAGCGGCTGCGGACTTTNC 303
QY 593 GAGCGCGGAGCGCTGGCGCTGATGAAGAGATCGAGCGCGCGGAGGCGCTGCGAG 652
Db 304 GAGCGCGGAG-GCCTNGGCGTGTGATGAAGAGATCGCGGCGCGGAGGCGGTNCAAN 362
QY 653 AGCGTCCAGAGGTGTTTCGGGCGCGGCTGCGGCGCGGCGGCGGAGCGGTGCGG 712
Db 363 AGGTGCAAGAGTNTTTTCGGGCGCGG-GTTCGCGGCGGCGGCGGCGGCGGAGCGTTGGG 421
QY 713 GAGCTGAGGTGCA 726
Db 422 GNTTNNAGTTNA 435

RESULT 6

US-08-799-173A-16
; Sequence 16, Application US/08799173A
; Patent No. 5871969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG,
; APPLICANT: PATRICK J. DILLON
; TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,173A
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-799-173A-16

Query Match 11.1%; Score 204.2; DB 2; Length 316;
Best Local Similarity 85.9%; Pred. No. 4.5e-34;
Matches 275; Conservative 0; Mismatches 38; Indels 7; Gaps 5;

QY 552 GGAAGAACAGTACGTACGTACAGCGGCTCGCGGACTTTTCGGAGCGCGGAGGCGCTGGG 611
Db 1 GGCANNCCAGTACGTCA-TAACGGGCTCGCGGACTTTTCGGGA-NGCGGCGAGGCGCTGGG 58

QY 612 CGCTGATGAAGGAGATCGAGCGCGCGGAGCGGCTGCAGAGCGTGCACGAGGTGTTTT 671
Db 59 CGCTGATGAAGGAGATCAAGCGCGCGGAGCGGCTGCAGAG-GTGACAGAGTGTGTTTT 117
QY 672 CGCGCGCCCGTCCCGAGCGGCGGCGAGCGTGCAGAGCTGGAGGTGCAGCGCA 731
Db 118 CGCGCGCCG-GTNNCCAGCNCACCGGCGAGCGTGGGCGAACTGGNAGGTGCAGCGCA 176
QY 732 GGCACCTCGTGTCTGTTTGTGGTGGCATCTGCCAGCCCGACTGTTTCGTGGGCG 791
Db 177 GGCACCTCGTGTCTGTTTGTGGTGGCATCTGCCAGCCCGACTGTTTCGTGGGCG 236
QY 792 TGGACAGCTTGGAC--TGTGCGAGCGGCGGCGGCTTGCGGCGGAAACAGGCGGCGCTGAGC 848
Db 237 TGGACAGCTTGGACCTGTGANAACGGGCGGCGCTTTNGCGNAGAACAGGCGGCTTGAGC 296
QY 849 TGTACCCCTACGACGCGGCGG 868
Db 297 TGTANCCCTACGACGCGGCG 316

RESULT 7

US-08-799-173A-17
; Sequence 17, Application US/08799173A
; Patent No. 5871969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG,
; APPLICANT: PATRICK J. DILLON
; TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,173A
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-799-173A-17

Query Match 11.1%; Score 204.2; DB 2; Length 316;
Best Local Similarity 85.9%; Pred. No. 4.5e-34;
Matches 275; Conservative 0; Mismatches 38; Indels 7; Gaps 5;

QY 552 GGAAGAACAGTACGTACGTACAGCGGCTCGCGGACTTTTCGGAGCGCGGAGGCGCTGGG 611
Db 1 GGCANNCCAGTACGTCA-TAACGGGCTCGCGGACTTTTCGGGA-NGCGGCGAGGCGCTGGG 58
QY 612 CGCTGATGAAGGAGATCGAGCGCGGCGGCGGCGTGCAGAGCGTGCACGAGGTGTTTT 671

Db 59 CGCTGATGAAGAGATCAAGAGCGGGGGAGCGCTGCAGAG- GTGCACAGAGGTCTTTT 117
QY 672 CGGCGCCCGCTCCCGACGCGACGGCGACAGCTGCGCGAGCTGGAGGTGCAGCGCA 731
Db 118 CGGCGCCCG- GTNNCCAGCGNACGCGACAGCTGCGGAACTGGNAGGTGCAGCGCA 176
QY 732 GGCACCTCGCTGCTCTGTTGTGTGCGCATCGTGCACAGCCCGACCTGGTTCGTGGGCG 791
Db 177 GGCACCTCGCTGCTCTGTTGTGTGCGCATCGTGCACAGCCCGACCTGGTTCGTGGGCG 236
QY 792 TGCACAGCTCGACC---TGTCGGAGGGGACCGTTCGCGGGAACAGCGCGCGCTGGACC 848
Db 237 TGCACAGCTCGGACCTGTGANAACGGGACCTTTTTCGNGNAACAGCGCGCGTTCGACC 296
QY 849 TGTACCCCTACGACGCGGG 868
Db 297 TGTANCCCTACGACGCGGG 316

RESULT 8

US-07-862-021B-9
; Sequence 9, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, AviHu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP U1

INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4029 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 226...2647

US-07-862-021B-9

Query Match 3.9%; Score 71.6; DB 1; Length 4029;
Best Local Similarity 50.2%; Pred. No. 2.2e-06;
Matches 269; Conservative 0; Mismatches 249; Indels 18; Gaps 3;

QY 400 CTGTTCCGCGCAGCCCGCGCCCAATACAGCATCACCTTCACGGCGCAAGTGGAGCCAGAC 459
Db 819 CTGTCGCGCTCGGGAACGTGCCAAGTACAGCTACGTTTATGGAAGTGGTTCGAGAA 878

QY 460 GGCCTTCCCAAGAGTACCCCTCTGTCGCGCCCTCGCAGTGTCTTCCTGCTCTGCTGG 519
Db 879 GACTCATCCAAAGGATTACCC-----TCGTCGGGCTAATCATCTGCTGCTCCATCATTTGG 932
QY 520 GGCCTCGGATAGCTCGGACTACAGCATGTGGAGGAGAACACAGTACGTACGTAAAGGGCT 579
Db 933 CGGATCCCACTCCAAAGAACTAGCTGCTGTGGGAGTACGGAGGGTATGCCAGTGAAGGGT 992
QY 580 GGCCTGCTTGGGAGCGCGGAGCGCTGCGCTGATGAAGAGAGATCGAGCGCGCGGG 639
Db 993 CAAGCAAGTTGCTGAACCTTGGCTCACCAGTAAATGAGGAGAGAAATTCGACACAGAG 1052
QY 640 GGAGCGCTGCAGAGCGTGCACAGGT-----GTTTCGCGCGCGCGCTCCCGAG 690
Db 1053 TGATGAAGTCTCTCACTGTCATCAAGCCAAAGCCAGTGGCCATCTGCGACGCTGTCAA 1112
QY 691 CGGACCGGGGAGAGCTGCGGCGAGCTGGAGTGCAGCGCAGGCACTCGCTGCTCTGTT 750
Db 1113 TGTGAGAGCAGCACCTTCAGCGCGAATCTCAGTGGAGCAGACACGCACTTGTATGCTCTT 1172
QY 751 TGTGCTGCGCATCGTGCACGCGCCGACTGTTCTGGGCGTGGAGCGCTGGACCTGGACTGTG 810
Db 1173 CCTAACATGATGGCGCCCGCTCTGACTGGAACGTGGCGCTATCTGCAGAGATCTGTG 1232
QY 811 CGACGGGAGCCTTGGCGGGAACAG---GCGCGCTGGACCTGTACCCCTACGACGCGGG 867
Db 1233 CACCAAGGAGTGTGCTGGTCCAGAAAGTGTGCAGGACCTAATTCCTGGGAGTGTGG 1292
QY 868 GACGGACAGCGCTTCACCTTCTCTCCCACTTCGCCACCATCCCGCAGGACA 933
Db 1293 CACGGACAGCGGGTGACCTAGAGTCAACAAAGCCCAATTCCTTCAGGAAA 1348

RESULT 9

US-08-313-288B-9
; Sequence 9, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and AviHu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 226...2647
; US-08-313-288B-9

Query Match
Best Local Similarity 3.9%; Score 71.6; DB 1; Length 4029;
Matches 269; Conservative 0; Mismatches 249; Indels 18; Gaps 3;

QY 400 CTGTTCCGCGCAGAGCCCGCCCAATACAGATCACCCTTCACGGGCAAGTGGAGCCAGAC 459
Db 819 CTGCTCGGCGCTGGGAACCTGCGCAAGTACAGACTACGTTTATGGGAACCTGGTCGGAGAA 878
QY 460 GCGCTTCCCAAGCAGTACCCCTGTTCGCGCCCGCCCTGCGCAGTGTCTTCGCTGCTGG 519
Db 879 GACTCATCCAAAGGATTACCC-----TCGTCGGGCTAATCACTGTGCTGCCATCATGG 932
QY 520 GCGCGCGCATAGCTCCGACTACAGATGTGGAGGAAGAACCACTAGTACGTACGTAAACGGGT 579
Db 933 CGGATCCCACTCCAAAGAACTACGTGTGTTGGAGTACGGAGGATGCCAGTGAAGGGGT 992
QY 580 GCGCGACTTTGCGGCGCGCGCGGAGCCCTGGCGCTGATCAAGGAGATCAGGCGCGCGG 639
Db 993 CAAGCAAGTGTGTAAGTGGCTTCACAGTAAATAATGGAGGAAGAAATTCGACACAGAG 1052
QY 640 GGAGGCGCTGCAGAGCGTGCACGAGT-----GTTTTCGGCGCCCGCCCTGCCAG 690
Db 1053 TGATGAAGTCTCCTACATCAAGCAAGCAAGCCAGTGGCCATCTCTGGCAGCTGTCAA 1112
QY 691 CGCAGCGGCGCAGAGCTCGCGGAGCTGGAGTGTGAGCGAGCAGTCTGCTGTCTGCTGG 750
Db 1113 TGTGAGAGCAGCACCCTCAGCCGAATCTCAGTGGAGCAGCAGCAGCTTGATGCTCT 1172
QY 751 TGTGTTGCGCATCGTCCCGCAGCCCGCAGTGTCTGTTGGCGTGGAGAGCTGCACCTGTG 810
Db 1173 CCTAACCATGATGGGCGCCAGTCTGACTGGAACCTGGGCGCTATCTGCAGAGGATCTGT 1232
QY 811 CGACGGGCGCAGTGTGGCGGGAACAG---GCGGCGCTGGAGCTGTACCCCTACGACCGCG 867
Db 1233 CACCAAGAGTGTGGCTGGGTCCAGAAAGTGTGCAGAGGCTAATTCCTCTGGATCTGTG 1292
QY 868 GACGGACAGCGGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 923
Db 1293 CACGGACAGCGGGGTGACCTACGAGTACCAAAAGCCCAAAATTCCTCAGGAAA 1348

RESULT 10
PCT-US93-03164-9
; Sequence 9, Application PC/TUS9303164.
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03164
; FILING DATE: 19930402
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```

;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4029 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 226...2647
; PCT-US93-03164-9

Query Match
Best Local Similarity 3.9%; Score 71.6; DB 5; Length 4029;
Matches 269; Conservative 0; Mismatches 249; Indels 18; Gaps 3;

QY 400 CTGTTCCGCGCAGAGCCCGCCCAATACAGATCACCCTTCACGGGCAAGTGGAGCCAGAC 459
Db 819 CTGCTCGGCGCTGGGAACCTGCGCAAGTACAGACTACGTTTATGGGAACCTGGTCGGAGAA 878
QY 460 GCGCTTCCCAAGCAGTACCCCTGTTCGCGCCCGCCCTGCGCAGTGTCTTCGCTGCTGG 519
Db 879 GACTCATCCAAAGGATTACCC-----TCGTCGGGCTAATCACTGTGCTGCCATCATGG 932
QY 520 GCGCGCGCATAGCTCCGACTACAGATGTGGAGGAAGAACCACTAGTACGTACGTAAACGGGT 579
Db 933 CGGATCCCACTCCAAAGAACTACGTGTGTTGGAGTACGGAGGATGCCAGTGAAGGGGT 992
QY 580 GCGCGACTTTGCGGCGCGCGGAGCCCTGGCGCTGATCAAGGAGATCAGGCGCGCGG 639
Db 993 CAAGCAAGTGTGTAAGTGGCTTCACAGTAAATAATGGAGGAAGAAATTCGACACAGAG 1052
QY 640 GGAGGCGCTGCAGAGCGTGCACGAGT-----GTTTTCGGCGCCCGCCCTGCCAG 690
Db 1053 TGATGAAGTCTCCTACATCAAGCAAGCAAGCCAGTGGCCATCTCTGGCAGCTGTCAA 1112
QY 691 CGCAGCGGCGCAGAGCTCGCGGAGCTGGAGTGTGAGCGAGCAGTCTGCTGTCTGCTGG 750
Db 1113 TGTGAGAGCAGCACCCTCAGCCGAATCTCAGTGGAGCAGCAGCAGCTTGATGCTCT 1172
QY 751 TGTGTTGCGCATCGTCCCGCAGCCCGCAGTGTCTGTTGGCGTGGAGAGCTGCACCTGTG 810
Db 1173 CCTAACCATGATGGGCGCCAGTCTGACTGGAACCTGGGCGCTATCTGCAGAGGATCTGT 1232
QY 811 CGACGGGCGCAGTGTGGCGGGAACAG---GCGGCGCTGGAGCTGTACCCCTACGACCGCG 867
Db 1233 CACCAAGAGTGTGGCTGGGTCCAGAAAGTGTGCAGAGGCTAATTCCTCTGGATCTGTG 1292
QY 868 GACGGACAGCGGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 923
Db 1293 CACGGACAGCGGGGTGACCTACGAGTACCAAAAGCCCAAAATTCCTCAGGAAA 1348

RESULT 11
US-07-862-021B-11
; Sequence 11, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```


Db 949 GATGAGTTTAAACAGTCATCAAGGCAAAAGCACAGTCGGCTGCCTGGCAGCCTCTGAAT 1008
QY 692 GGCACCGGCGAGACGTCGGCGGAGCTGGAGTGCACGCGCAGGCACTCGTGGTCTGTTT 751
Db 1009 GTGAGAGCTGCTCCCTCTCTGAGTTTCTGTGTATCGCGCACCGGACCTGATGTCCTTC 1068
QY 752 GTGGTGGCATCGTGGCCAGCCCGGACTGGTTCTGGTGGCGGTGGACAGCTGGACCTGTGC 811
Db 1069 CTCACCATGCTGGGGCCAGTCCGACTGGAATGTGGGCTGTCTGCTGAGGACCTCTGC 1128
QY 812 GACGGGAGCGTTGGGGGGAACAG---GCGGCGCTGGAGCTGTACCCCTACGACGCCGG 868
Db 1129 ACCAAGGACTGTGGTGGGTTTCAGAAAGTCGTGCAGGATTTAATCCCTGGGATGCGGC 1188
QY 869 AGCGACGCGGCTTACCTTCTCCCTCCCACTTCGCGCACCATCCCGCAGGACGGTG 928
Db 1189 ACAGACAGTGGGCTACCTATGAGTCACCCCAACACCTACAGTTCTCTCAAGAGAAGATT 1248
QY 929 A 929
Db 1249 A 1249

RESULT 13

PCT-US93-03164-11
; Sequence 11, Application PC/TUS9303164
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03164
; FILING DATE: 19930402
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3226 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136...2543
PCT-US93-03164-11

Query Match 3.4%; Score 62.2; DB 5; Length 3226;

Best Local Similarity 49.0%; Pred. No. 0.00019;
Matches 265; Conservative 0; Mismatches 258; Indels 18; Gaps 3;
QY 401 TGTTCGGCCAGAGCCCGCGCAAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGACG 460
Db 715 TCTGTGCTGTGGAACTCGCAAAATACAGGCTAACGTTTTATGGAATTTGGTCGGAATAA 774
QY 461 GCCTTCCCAAGCAGTACCCCTGTTCGCCGCCCTGCGCAGTGGTCTTCGCTGCTGGGG 520
Db 775 ACACATCCCAAGACATTTC-----TCGGGCGACCAACCATTTGGTCTCGGATCATGGT 828
QY 521 GCGGCGCATAGTCCGATACAGCATGTGGAGGAAAGAACACAGTACGTACGTACACGGCTG 580
Db 829 AGCTCTCACTCAAGAACTACATCCCTTTGGAGTATGGAGGGTATGCTAGTGAAGGTGTC 888
QY 581 CCGGACTTTTGGGAGCGCGGAGGCTGGGCGCTGATGAAGGAGATCGAGGCGCGGG 640
Db 889 AAGCAGGTTTGCAGAGCTGGGATCCCCAGTCAAGATGGAAGAAGAAATTCGACAACAAAGT 948
QY 641 GAGGCGCTCAGAGCGTCGACGAGG-----TGTTCGGCGCGCGCGCTCCCGCAGC 691
Db 949 GATGAGTTTAAACAGTCATCAAGGCAAAAGCACAGTGGCTGCCTGGCAGCCTCTGAAT 1008
QY 692 GGCACCGGCGAGACGTCGGCGGAGTGGAGTGCACGCGCAGGCACTCGTGGTCTGTTT 751
Db 1009 GTGAGAGCTGCTCCCTCTCTGAGTTTCTGTGTATCGCGCACCGGACCTGATGTCCTTC 1068
QY 752 GTGGTGGCATCGTGGCCAGCCCGGACTGGTTCTGGTGGCGGTGGACAGCTTGGACCTGTGC 811
Db 1069 CTCACCATGCTGGGGCCAGTCCGACTGGAATGTGGGCTGTCTGCTGAGGACCTCTGC 1128
QY 812 GACGGGAGCGTTGGCGGGAACAG---GCGGCGCTGGAGCTGTACCCCTACGACGCCGG 868
Db 1129 ACCAAGGACTGTGGTGGGTTTCAGAAAGTCGTGCAGGATTTAATCCCTGGGATGCGGC 1188
QY 869 ACAGACAGCGCTTACCTTCTCCCTCCCACTTCGCGCACCATCCCGCAGGACAGCGTG 928
Db 1189 ACAGACAGTGGGCTACCTATGAGTCACCCCAACACCTACAGTTCTCTCAAGAGAAGATT 1248
QY 929 A 929
Db 1249 A 1249

RESULT 14
US-08-818-112-12
; Sequence 12, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112

```
;
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-818-112-12

Query Match 3.1%; Score 57.6; DB 4; Length 1155;
Best Local Similarity 43.9%; Pred. No. 0.0014;
Matches 246; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

QY 335 GCCTCTGGCTCTCTCTCTGGCACTCTCGGCGCGCGCGCGAGCTCTTGGGGAGAG 394
Db 141 GCCACCGGAGCTACCCGGGTCAAGGTACCGGTGCGGGTGTCTCAAGCGGTGGCC 200
QY 395 TCCATCTGTTCCGCCAGAGCCCGGCAATACAGCATCACCTTACGGGCAAGTGGAGC 454
Db 201 GCGAGCGCGCATCCACACCGAGTTCAACAGACGATCGCGGACCGCGGCTCTCC 260
QY 455 CAGACGGCTTCCCAAGCAGTACCCCTGTTCGGCGCCCGCTGCGCAGTGGTTCGCTG 514
Db 261 TCGACGAACCACTTATTCGTTGTCGCGTCCCGGGTGTTCGGTGTGCGGATCCGG 320
QY 515 CTGGGGCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTACGTAAC 574
Db 321 GACGCTTCTGTTGCGGGACGGTTACGCCGAGTTGATGACGGCGGCTACAGCCAGATC 380
QY 575 GGGCTGCGCGACTTTTCGGAGCGCGCGAGGCTTGGCGCTGATGAAGGAGATCGAGCG 634
Db 381 GAGATCGGATCATGGGCACTGTCCGGGAGGAGGCTCATCGAGCGTTCAACACC 440
QY 635 GCGGGGAGCGCTGCAGAGCGTGCAGAGGTGTTTCGGCGCCCGCTCCCGAGCGGC 694
Db 441 GGGGAGGACCTGATTTCTGTTCTCGGTCGCCGGTGTTCGGTGTGCGGATCGAGGTC 500
QY 695 ACCGGGCGAGCTCGGGGAGCTGGAGGTGCGAGCGAGGCTCGCTGCTCTGTTTGTG 754
Db 501 ACCGGGAGTTTGGGGCGCGGTCAGGCGATGTCTACGGGCTGTTTACGGGTTGAGC 560
QY 755 GTGCGCATCTGCCAGCCCGGAGCTGGTGTGGGGCTGGAGCGCTGAGACCTGTGCGAC 814
Db 561 GCCTACGGCTGTGCGAGCAGTTGAAATCTCCACCGAGGAAGCCAAACGAGCAGATGAC 620
QY 815 GGGGACCGTTGGGGGAACAGCGGGCGCTGGACCTGTACCCCTACGACCGCGGAGGAC 874
Db 621 GCGATTTTCGCGGATTCGGGGGTTGCGGAGTACCTCGCGCGCTAGTCGAGCGGGCC 680
QY 875 AGCGGCTTCACCTTCCTC 894
Db 681 CGCAGAGCGGCTACACCTC 700

RESULT 15
US-08-818-111-12
; Sequence 12, Application US/08818111
; Patent No. 633852
;
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, David C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
```


QY 875 AGCGGCTTCACCTTCCTC 894
 || | | | |
Db 681 CGCAGGACGGCTACACCTC 700

Search completed: October 27, 2002, 11:16:08
Job time : 101 secs

